# PCT

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| | | |
|---|---|---|
| **(51) International Patent Classification 5 :**<br><br>C07K 5/00, 7/00, C12Q 1/68, C08L 77/00, A61K 48/00 | **A1** | **(11) International Publication Number:** WO 95/01370<br><br>**(43) International Publication Date:** 12 January 1995 (12.01.95) |

**(21) International Application Number:** PCT/US94/07319

**(22) International Filing Date:** 28 June 1994 (28.06.94)

**(30) Priority Data:**
08/088,658      2 July 1993 (02.07.93)      US

**(71) Applicant:** ISIS PHARMACEUTICALS, INC. [US/US]; 2280 Faraday Avenue, Carlsbad, CA 92008 (US).

**(71)(72) Applicants and Inventors:** BUCHARDT, Ole [DK/DK]; Sondergardsvej 73, DK-3500 Vaerlose (DK). EGHOLM, Michael [DK/DK]; Johnstrup Allee 3, DK-1923 Fredriksberg (DK). NIELSEN, Peter, E. [DK/DK]; Hjortevanget 509, DK-2980 Koddedal (DK). BERG, Rolf, H. [DK/DK]; Strandvaenget 6, DK-2960 Rungsted Kyst (DK).

**(72) Inventors:** ECKER, David, J.; 1041 Saxony Road, Leucadia, CA 92024 (US). MOLLEGAARD, Niels, E.; Benvedvaenget 27, Sorgenfri, DK-2830 Virum (DK).

**(74) Agents:** CALDWELL, John, W. et al.; Woodcock Washburn Kurtz Mackiewicz & Norris, 46th floor, One Liberty PLace, Philadelphia, PA 19103 (US).

**(81) Designated States:** CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

**Published**
*With international search report.*

---

**(54) Title:** HIGHER ORDER STRUCTURE AND BINDING OF PEPTIDE NUCLEIC ACIDS

**(57) Abstract**

Peptide nucleic acids and analogues of peptide nucleic acids are used to form duplex, triplex, and other structures with nucleic acids and to modify nucleic acids. The peptide nucleic acids and analogues thereof also are used to modulate protein activity through, for example, transcription arrest, transcription initiation, and site specific cleavage of nucleic acids.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| AT | Austria | GB | United Kingdom | MR | Mauritania |
| AU | Australia | GE | Georgia | MW | Malawi |
| BB | Barbados | GN | Guinea | NE | Niger |
| BE | Belgium | GR | Greece | NL | Netherlands |
| BF | Burkina Faso | HU | Hungary | NO | Norway |
| BG | Bulgaria | IE | Ireland | NZ | New Zealand |
| BJ | Benin | IT | Italy | PL | Poland |
| BR | Brazil | JP | Japan | PT | Portugal |
| BY | Belarus | KE | Kenya | RO | Romania |
| CA | Canada | KG | Kyrgystan | RU | Russian Federation |
| CF | Central African Republic | KP | Democratic People's Republic | SD | Sudan |
| CG | Congo | | of Korea | SE | Sweden |
| CH | Switzerland | KR | Republic of Korea | SI | Slovenia |
| CI | Côte d'Ivoire | KZ | Kazakhstan | SK | Slovakia |
| CM | Cameroon | LI | Liechtenstein | SN | Senegal |
| CN | China | LK | Sri Lanka | TD | Chad |
| CS | Czechoslovakia | LU | Luxembourg | TG | Togo |
| CZ | Czech Republic | LV | Latvia | TJ | Tajikistan |
| DE | Germany | MC | Monaco | TT | Trinidad and Tobago |
| DK | Denmark | MD | Republic of Moldova | UA | Ukraine |
| ES | Spain | MG | Madagascar | US | United States of America |
| FI | Finland | ML | Mali | UZ | Uzbekistan |
| FR | France | MN | Mongolia | VN | Viet Nam |
| GA | Gabon | | | | |

## HIGHER ORDER STRUCTURE AND BINDING OF PEPTIDE NUCLEIC ACIDS

### RELATED APPLICATION

This patent application is a continuation-in-part of
5   patent application Serial No. 08/054,363, filed April 26, 1993,
which is a continuation-in-part of application PCT EP92/01219,
filed May 19, 1992 and published November 26, 1992 as WO
92/20702. The entire contents of each of the foregoing patent
applications are incorporated herein by reference.

10  ### FIELD OF THE INVENTION

This invention is directed to compounds that form triple-
stranded structures with single-stranded and double-stranded
nucleic acids. It is further directed to the use of such
compounds to cause strand displacement in double-stranded
15  nucleic acids. The invention further is directed to processes
for modifying double-stranded nucleic acid utilizing such
strand displacement. Such processes for modifying double-
stranded nucleic acids include cleavage of the nucleic acid
strand or strands. In particular, such cleavage includes
20  sequence specific cleavage of double-stranded nucleic acids
using a nuclease which normally is nonsequence specific. Such
processes also include transcription inhibition or arrest as
well as transcription initiation. The processes of the
invention are effected, in particular, with compounds that
25  include naturally-occurring nucleobases or other nucleobase-
binding moieties covalently bound to a polyamide backbone.

### BACKGROUND OF THE INVENTION

The function of a gene starts by transcription of its
information to a messenger RNA (mRNA) which, by interaction
30  with the ribosomal complex, directs the synthesis of a protein

- 2 -

## BACKGROUND OF THE INVENTION

The function of a gene starts by transcription of its information to a messenger RNA (mRNA) which, by interaction with the ribosomal complex, directs the synthesis of a protein

5   coded for by the mRNA sequence.  The synthetic process is known as translation.  Translation requires the presence of various co-factors and building blocks, the amino acids, and their transfer RNAs (tRNA), all of which are present in normal cells.

Transcription initiation requires specific recognition

10  of a promoter DNA sequence by the RNA-synthesizing enzyme, RNA polymerase.  In many cases in prokaryotic cells, and probably in all cases in eukaryotic cells, this recognition is preceded by sequence-specific binding of a protein transcription factor to the promoter.  Other proteins which bind to the promoter,

15  but whose binding prohibits action of RNA polymerase, are known as repressors.  Thus, gene activation typically is regulated positively    by    transcription    factors    and    negatively    by repressors.

Most conventional drugs function by interaction with and

20  modulation of one or more targeted endogenous proteins, e.g., enzymes.    Typical   daily   doses   of   drugs   are   from   $10^{-5}$-$10^{-1}$ millimoles per kilogram of body weight or $10^{-3}$-10 millimoles for a 100 kilogram person.  If this modulation instead could be effected  by  interaction  with  and  inactivation  of  mRNA,  a

25  dramatic reduction in the necessary amount of drug necessary likely could be achieved, along with a corresponding reduction in side effects.  Further reductions could be effected if such interaction  could  be  rendered  site-specific.    Given  that  a functioning gene continually produces mRNA, it would thus be

30  even more advantageous if gene transcription could be arrested in its entirety.

Synthetic  reagents  that  bind  sequence  selectively  to single-stranded and especially to double-stranded nucleic acids are   of   great   interest   in   molecular   biology   and

35  medicinal/chemistry, since such reagents may provide the tools for developing gene-targeted drugs and other sequence-specific

gene modulators.  Until now oligonucleotides and their close
analogues have presented the best candidates for such reagents.

        Oligodeoxyribonucleotides as long as 100 base pairs (bp)
are  routinely  synthesized  by  solid  phase  methods  using

5  commercially  available,  fully  automatic  synthesis  machines.
Oligoribonucleotides,  however,  are  much  less  stable  than
oligodeoxyribonucleotides, a fact which has contributed to the
more prevalent use of oligodeoxyribonucleotides in medical and
biological research directed to, for example, gene therapy and

10  the  regulation  of  transcription  and  translation.   Synthetic
oligodeoxynucleotides  are  being  investigated  for  used  as
antisense probes to block and eventually breakdown mRNA.

        It also may be possible to modulate the genome of an
animal  by,  for  example,  triple  helix  formation  using

15  oligonucleotides  or  other  DNA  recognizing  agents.   However,
there are a number of drawbacks associated with oligonucleotide
triple  helix  formation.   For  example,  triple  helix  formation
generally has only been obtained using homopurine sequences and
requires  unphysiologically  high  ionic  strength  and  low  pH.

20  Whether used as antisense reagents or a triplexing structures,
unmodified oligonucleotides are unpractical because they have
short in vivo half-lives.  To circumvent this, oligonucleotide
analogues have been used.

        These  areas  for  concern  have  resulted  in  an  extensive

25  search for improvements and alternatives.  For example, the
problems arising in connection with double-stranded DNA (dsDNA)
recognition through triple helix formation have been diminished
by a clever "switch back" chemical linking whereby a sequence
of polypurine on one strand is recognized, and by "switching

30  back",  a  homopurine  sequence  on  the  other  strand  can  be
recognized.  See,  e.g.,  McCurdy,  Moulds,  and  Froehler,
Nucleosides, in press.  Also, helix formation has been obtained
by using artificial bases, thereby improving binding conditions
with regard to ionic strength and pH.

35        In  order  to  improve  half-life  as  well  as  membrane
penetration,  a  large  number  of  variations  in  polynucleotide
backbones has been undertaken.  These variations include the

- 4 -

use of methylphosphonates, monothiophosphates, dithiophos-
phates, phosphoramidates, phosphate esters, bridged
phosphoramidates, bridged phosphorothioates, bridged
methylenephosphonates, dephospho internucleotide analogs with
5   siloxane bridges, carbonate bridges, carboxymethyl ester
bridges, acetamide bridges, carbamate bridges, thioether,
sulfoxy, sulfono bridges, various "plastic" DNAs, α-anomeric
bridges, and borane derivatives. The great majority of these
backbone modifications have decreased the stability of hybrids
10  formed between a modified oligonucleotide and its complementary
native oligonucleotide, as assayed by measuring $T_m$ values.
Consequently, it is generally believed in the art that backbone
modifications destabilize such hybrids, i.e., result in lower
$T_m$ values, and should be kept to a minimum.

15      The discovery of sequence specific endonucleases
(restriction enzymes) was an essential step in the development
of biotechnology, enabling DNA to be cut at precisely specified
locations containing specific base sequences. However, although
the range of restriction enzymes now known is extensive, there
20  is still a need to obtain greater flexibility in the ability to
recognize particular sequences in double-stranded nucleic acids
and to cleave the nucleic acid specifically at or about the
recognized sequence.

        Most restriction enzymes recognize quartet or sextet DNA
25  sequences and only a very few require octets for recognition.
However, restriction enzymes have been identified and isolated
only for a small subset of all possible sequences within these
constraints. A need exists, especially in connection with the
study of large genomic DNA molecules, in general, and with the
30  human genome project, in particular, to recognize and
specifically cleave DNA molecules at more rarely occurring
sites, e.g., sites defined by about fifteen base pairs.

        Efforts have therefore been made to create artificial
"restriction enzymes" or to modify the procedures for using
35  existing restriction enzymes for this purpose. Methods
investigated include the development of oligonucleotides
capable of binding sequence specifically via triple helix

- 5 -

formation to double-stranded DNA tagged with chemical groups (e.g., photochemical groups) able to cleave DNA or with non-specific DNA cleaving enzymes and other such modifications consistent with an "Achilles heel" general strategy.    Such
5   methods are described by: Francois, J.C., et al. PNAS 86,9702-9706 (1989); Perrouault, L., et al Nature 344,358-360 (1990); Strobel, S.A. & Dervan P.B. Science 249,73-75 (1990); Pei, D., Corey D.R. & Schultz P.G. PNAS 87,9858 (1990); Beal, P.A. & Dervan P.B., Science 251,1360 (1991); Hanvey, J.C., Shimizu M.
10  & Wells R.D. NAR 18,157-161 (1990); Koob, H. & Szybalski W. Science 250,271 (1990); Strobel, S.A. & Dervan P.B. Nature #50,172 (1991); and Ferrin, L.J. & Camerini-Otero R.D. Science 254,1494-1497 (1991).

In    Patent    Cooperation    Treaty    Applications    No.
15  PCT/EP92/01220, and PCT/EP92/01219, both filed on 22nd May 1992, we described certain nucleic acid analogue compounds that have a strong sequence specific DNA binding ability.  Examples of such compounds were also disclosed by us in *Science* **1991**, *254* 1497-1500.  We have shown that a nucleic acid analogue of
20  this type containing 10 bases hybridized to a non-terminal region of a double-stranded DNA and rendered the strand of DNA which was non-complementary to the nucleic acid analogue susceptible to degradation by $S_1$ nuclease.    No increased cleavage of the DNA strand complementary to the nucleic acid
25  analogue was seen, so no cleavage of double-stranded DNA was obtained.

**OBJECTS OF THE INVENTION**

It is an object of the invention to provide compounds that bind DNA and RNA strands.
30      It is a further object of the invention to provide triplex structures between DNA or RNA strands and these compounds.

It is yet another object to provide compounds other than RNA that can bind one strand of a double-stranded
35  polynucleotide, thereby displacing the other strand.

- 6 -

It is still another object to provide therapeutic, diagnostic, and prophylactic methods that employ such compounds.

**BRIEF DESCRIPTION OF THE INVENTION**

5    In the cell, DNA exists as a double stranded structure. During certain cellular events, as for instance transcription or during cell division, portions of the double stranded DNA are transiently denatured to single strand. Further DNA can be isolated outside of a cell and can be purposefully denatured to

10   single stranded DNA. RNA generally exist as a single stranded structure; however, in a local area of secondary structure a RNA, as for instance the stem of a stem loop structure, the RNA can exist as a double stranded structure.

We have found that certain compounds that have

15   nucleobases attached to an aminoethylglycine backbone and other like backbones including polyamides, polythioamides, polysulfinamides and polysulfonamides, which compounds we call peptide nucleic acids or PNA, surprisingly bind strongly and sequence selectively to both RNA and DNA.

20   We have surprisingly found that these PNA compounds recognize and bind sequence-selectively and strand-selectively to double-stranded DNA (dsDNA). We have found that the binding to double-stranded DNA is accomplished via strand displacement, in which the PNA binds via Watson-Crick binding to its

25   complementary strand and extrudes the other strand in a virtually single-stranded conformation. We have also surprisingly found that these PNA compounds recognize and bind sequence-selectively to single-stranded DNA (ssDNA) and to RNA.

The recognition of PNA to RNA, ssDNA or dsDNA can take

30   place in sequences at least 5 bases long. A more preferred recognition sequence length is 5-60 base pairs long. Sequences between 10 and 20 bases are of particular interest since this is the range within which unique DNA sequences of prokaryotes and eukaryotes are found. Sequences of 17-18 bases are of

35   special interest since this is the length of unique sequences in the human genome.

- 7 -

We have further surprisingly found that the PNA compounds are able to form triple helices with dsDNA. We have found that PNA compounds are able to form triple helices with RNA and ssDNA. The resulting triplexes, e.g., $(PNA)_2$/DNA or $(PNA)_2$/RNA,

5  surprisingly have very high thermal stability. It has been found that the PNA binds with a DNA or RNA in either orientation, i.e., the antiparallel orientation where the amino-terminal of the PNA faces the 3' end of the nucleic acid or the parallel orientation where the amino-terminal of the PNA

10 faces the 5' end of the nucleic acid. PNAs are able to form triple helices wherein a first PNA strand binds with RNA or ssDNA and a second PNA strand binds with the resulting double helix or with the first PNA strand.

We further have found that the PNA compounds are able to

15 form triple helices wherein a first PNA strand binds with the ssDNA or RNA or to one of the strands of dsDNA and in doing so displaces the other strand, and a second PNA strand then binds with the resulting double helix. While we do not wish to be bound by theory, it is further believed that other triple

20 helices might be formed wherein a single PNA strand binds to two single stranded nucleic acids strands. In binding with nucleic acids both Watson-Crick and Hoogsteen bind is utilized. It is further believed that PNA might also bind via reverse Hoogsteen binding.

25     We have further surprisingly found that the PNA compounds form double helices with RNA and ssDNA. Such double helices are hetero duplex structures between the PNA and the respective nucleic acid. Such double helices are preferably helices formed when the PNA strand includes a mixture of both

30 pyrimidine and purines nucleobases.

For therapeutic use of PNA compounds the targets of the PNA compounds would generally be double stranded DNA and RNA. For diagnostic use, investigations methods and reagents where DNA is isolated outside of a cell, the DNA can be denatured to

35 single stranded DNA and use of the PNA compound would be targeted to such single stranded DNA as well as RNA.

- 8 -

PNA compounds useful to effect binding to RNA, ssDNA and dsDNA and to form duplex and triplex complexes are in one sense polymeric strands formed from a polyamide, polythioamide, polysulfinamide or polysulfonamide backbone with a plurality of
5  ligands located at spaced locations along the backbone. At least some of the ligands are capable of hydrogen bonding with other ligands either on the compounds or nucleic acid ligands.

More preferred PNA compounds according to the invention have the formula:

10

$$Q\diagdown_{C^1}\diagup^{B^1}\diagdown_{D^1}\diagup^{G^1}\diagdown_{C^2}\diagup^{B^2}\diagdown_{D^2}\diagup^{G^2}\text{-------}_{C^n}\diagup^{B^n}\diagdown_{D^n}\diagdown^{I}$$

15
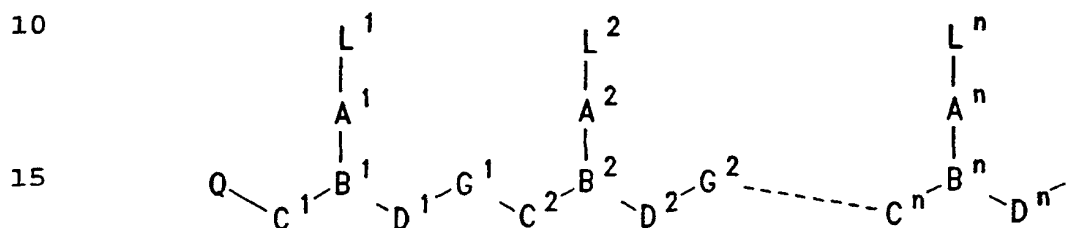
wherein:

20     n is at least 2,

each of $L^1$-$L^n$ is independently selected from the group consisting of hydrogen, hydroxy, $(C_1$-$C_4)$alkanoyl, naturally occurring nucleobases, non-naturally occurring nucleobases, aromatic moieties, DNA intercalators, nucleobase-binding
25  groups, heterocyclic moieties, and reporter ligands;

each of $C^1$-$C^n$ is $(CR^6R^7)_y$ where $R^6$ is hydrogen and $R^7$ is selected from the group consisting of the side chains of naturally occurring alpha amino acids, or $R^6$ and $R^7$ are independently selected from the group consisting of hydrogen,
30  $(C_2$-$C_6)$alkyl, aryl, aralkyl, heteroaryl, hydroxy, $(C_1$-$C_6)$alkoxy, $(C_1$-$C_6)$alkylthio, $NR^3R^4$ and $SR^5$, where $R^3$ and $R^4$ are as defined above, and $R^5$ is hydrogen, $(C_1$-$C_6)$alkyl, hydroxy-, alkoxy-, or alkylthio- substituted $(C_1$-$C_6)$alkyl, or $R^6$ and $R^7$ taken together complete an alicyclic or heterocyclic system;
35     each of $D^1$-$D^n$ is $(CR^6R^7)_z$ where $R^6$ and $R^7$ are as defined above;

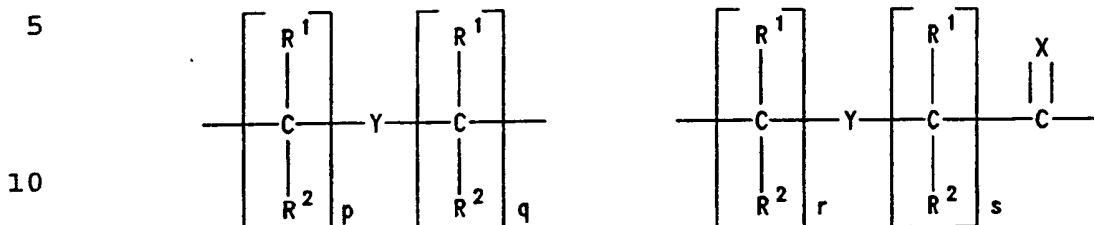each of y and z is zero or an integer from 1 to 10, the sum y + z being greater than 2 but not more than 10;

each of $G^1$-$G^{n-1}$ is $-NR^3CO-$, $-NR^3CS-$, $-NR^3SO-$ or $-NR^3SO_2-$, in
40  either orientation, where $R^3$ is as defined above;
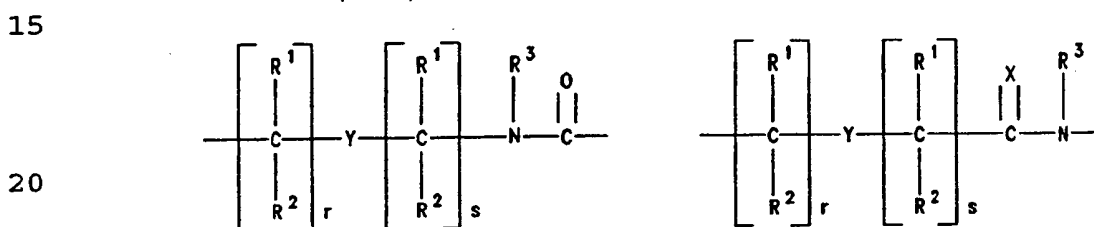
each of $A^1-A^n$ and $B^1-B^n$ are selected such that:

(a)    A is a group of formula (IIa), (IIb), (IIc) or (IId), and B is N or $R^3N^+$; or

(b)    A is a group of formula (IId) and B is CH;



(IIa)                                    (IIb)



(IIc)                                    (IId)

where:

X is O, S, Se, $NR^3$, $CH_2$ or $C(CH_3)_2$;

Y is a single bond, O, S or $NR^4$;

each of p and q is zero or an integer from 1 to 5, the sum p+q being not more than 10;

each of r and s is zero or an integer from 1 to 5, the sum r+s being not more than 10;

each $R^1$ and $R^2$ is independently selected from the group consisting of hydrogen, $(C_1-C_4)$alkyl which may be hydroxy- or alkoxy- or alkylthio-substituted, hydroxy, alkoxy, alkylthio, amino and halogen; and

each $R^3$ and $R^4$ is independently selected from the group consisting of hydrogen, $(C_1-C_4)$alkyl, hydroxy- or alkoxy- or alkylthio-substituted $(C_1-C_4)$alkyl, hydroxy, alkoxy, alkylthio and amino;

Q is $-CO_2H$, $-CONR'R''$, $-SO_3H$ or $-SO_2NR'R''$ or an activated derivative of $-CO_2H$ or $-SO_3H$; and

I is $-NHR'''R''''$ or $-NR'''C(O)R''''$, where R', R", R''' and R'''' are independently selected from the group consisting of hydrogen, alkyl, amino protecting groups, reporter ligands,

- 10 -

intercalators, chelators, peptides, proteins, carbohydrates, lipids, steroids, nucleosides, nucleotides, nucleotide diphosphates, nucleotide triphosphates, oligonucleotides, oligonucleosides and soluble and non-soluble polymers.

5        In the above structures wherein R', R", R''' and R'''' are oligonucleotides or oligonucleosides, such structures can be considered chimeric structures between PNA compounds and the oligonucleotide or oligonucleoside. Preferred PNA-containing compounds useful to effect binding to RNA, ssDNA and dsDNA and

10    to form triplexing structure are compounds of the formula III, IV or V:
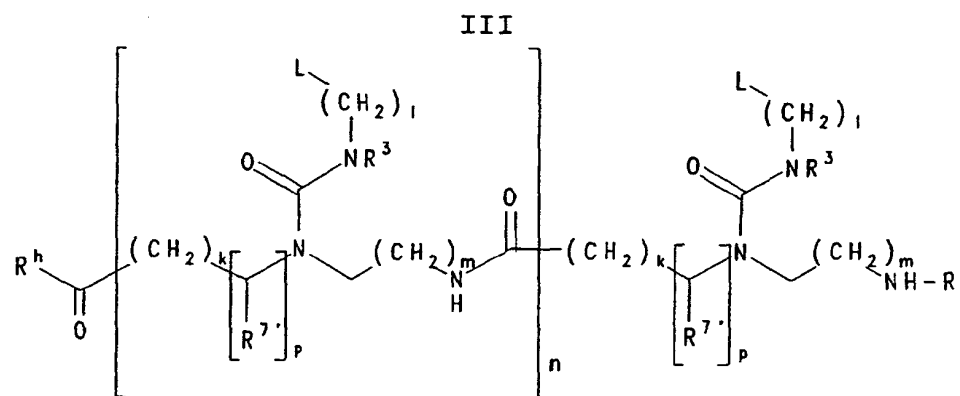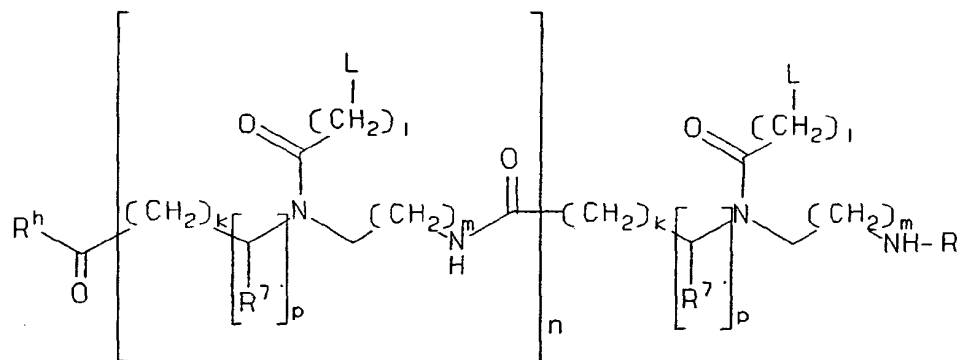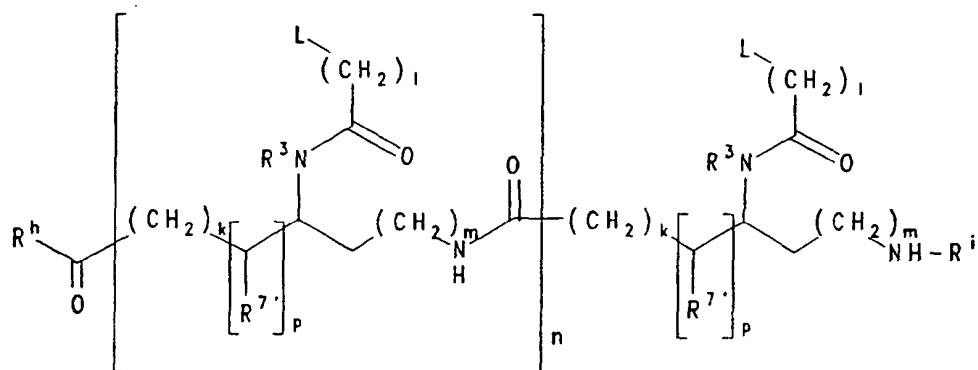


III



IV

15

20

V

wherein:

each L is independently selected from the group
consisting of hydrogen, phenyl, heterocyclic moieties,
naturally occurring nucleobases, and non-naturally occurring
nucleobases;

each $R^{7''}$ is independently selected from the group
consisting of hydrogen and the side chains of naturally
occurring alpha amino acids;

n is an integer greater than 1,

each k, l, and m is, independently, zero or an integer
from 1 to 5;

each p is zero or 1;

$R^h$ is OH, $NH_2$ or -$NHLysNH_2$; and

$R^i$ is H or $COCH_3$.

The improved binding of the PNA compounds of the
invention with single-stranded RNA and DNA renders them useful
as antisense agents. In addition, the binding to double-
stranded DNA renders these compounds useful for gene inhibition
via various mechanisms. Further, the binding to double-
stranded DNA renders these compounds useful as gene activators
to initiate transcription.

In one embodiment, the present invention provides methods
for inhibiting the expression of particular genes in the cells
of an organism, comprising administering to said organism a
reagent as defined above which binds specifically to sequences
of said genes.

- 12 -

In a further embodiment, the invention provides methods
for inhibiting transcription and/or replication of particular
genes or for modifying double-stranded DNA as, for instance, by
inducing degradation of particular regions of double-stranded
5   DNA in cells of an organism comprising administering to said
organism a reagent as defined above.

In a still further embodiment, the invention provides
methods for killing cells or virus by contacting said cells or
virus with a reagent as defined above which binds specifically
10  to sequences of the genome of said cells or virus.

A novel strategy for sequence-selective cleavage of
double-stranded DNA is described.  For cases were two closely
positioned homo-pyrimidine stretches (of 7-10 bases and
preferably on opposite strands) can be identified, this can be
15  done by synthesizing pairs of PNAs complementary to two parts
of this DNA sequence and separated by several base pairs.
These PNA molecules are then reacted with the dsDNA and the
resulting complex is allowed to react with an endonuclease.

In practicing certain embodiments of the invention, the
20  PNA compounds are able to recognize duplex DNA by displacing
one strand, thereby presumably generating a hetero duplex with
the other one.  Such recognition can take place with dsDNA
sequences 5-60 base pairs long.  Sequences between 10 and 20
bases are of interest since this is the range within which
25  unique DNA sequences of prokaryotes and eukaryotes are found.
Reagents which recognize 17-18 bases are of particular interest
since this is the length of unique sequences in the human
genome.

The PNA compounds are able to form triple helices with
30  dsDNA, ssDNA or RNA and double helices with RNA or ssDNA.  In
one embodiment of the invention, the PNA compounds form triple
helices wherein a first PNA strand binds with a nucleic acid
strand forming a hetero duplex and a second PNA strand then
binds with the resulting hetero duplex.  In other embodiments
35  of the invention, a PNA compound or a PNA chimera compound
forms triple helices wherein a single PNA strand or PNA chimera
strand binds with two nucleic acid strands, with a nucleic acid

strand and a PNA chimera strand or with two chimera PNA strands.

The invention further provides methods for inhibiting the action of restriction enzymes at restriction sites in nucleic
5  acids. Such methods comprise contacting a nucleic acid with a reagent as defined above under conditions effective to bind such reagent to the nucleic acid proximal to a restriction site.

The invention further provides methods of sequencing DNA
10  by binding the DNA with a reagent as defined above at a site proximal to a restriction site, cleaving the DNA with a restriction enzyme, and identifying the cleaved products.

The invention further provides methods for initiating transcription in cells or organisms comprising administering to
15  the organism a reagent as defined above which initiates transcription of a gene in such cells or organisms.

The invention further provides methods for modulating binding of RNA polymerase to dsDNA by binding the dsDNA with a reagent as defined above that binds with the DNA and then
20  exposing the complex formed thereby to a RNA polymerase.

The invention further provides methods for initiating transcription of a gene by binding the gene with a reagent as defined above that interacts with the gene to melt the double-stranded DNA of the gene and to form a transcription elongation
25  loop.

The invention further provides methods for binding RNA polymerase to dsDNA by contacting the dsDNA with a reagent as defined above that is capable of interacting with the DNA and then exposing the complex formed thereby to a RNA polymerase.
30  More particularly, the interaction is via binding to said dsDNA.

The invention further provides a hybrid complex for modulating transcription wherein the complex comprises dsDNA and a reagent as defined above that binds with the dsDNA, and
35  a RNA transferase.

The invention further provides a synthetic transcription factor comprising a dsDNA and a reagent as defined above that

- 14 -

is capable of interaction with the dsDNA.  More particularly, the interaction is via binding to said dsDNA.

The invention further provides specific gene activators comprising first and second strands, as defined above, that

5   have specific sequences that bind to selected DNA regions of the gene.

The invention further provides chimeric structures comprising PNAs and DNA or RNA.  Such chimeric structures will be used in place of or in addition to a normal PNA strand to

10  effect duplexing, triplexing, nucleic acid binding or protein binding.

**BRIEF DESCRIPTION OF THE DRAWINGS**

Figure 1 is an ethidium bromide stained gel showing

15  cleavage in a sequence specific manner by $S_1$ nuclease and a nucleic acid analogue of a double-stranded DNA having a single site for hybridization to the nucleic acid analogue (Example 1);

Figure 2 is an ethidium bromide stained gel showing a

20  similar cleavage of a double-stranded DNA having two closely spaced sites for hybridization to the nucleic acid analogue on the same DNA strand (Example 2);

Figure 3 is an ethidium bromide stained gel showing a similar cleavage of a double-stranded DNA having two closely

25  spaced sites for hybridization to a nucleic acid analogue on opposite strands (Example 3).

Figure 4 shows a PAGE autoradiograph demonstrating that PNAs-$T_{10}$, -$T_9C$ and -$T_8C_2$ bind to double-stranded DNA with high sequence specificy;

30  Figure 5 shows an electrophoretic gel staining demonstrating that restriction enzyme activity towards DNA is inhibited when PNA is bound proximal to the restriction enzyme recognition site;

Figure 6 shows a graph based on densitometric scanning

35  of PAGE autoradiographs demonstrating the kinetics of the binding of PNA-$T_{10}$ to a double-stranded target;

- 15 -

Figure 7 shows a graph based on densitometric scanning of PAGE autoradiographs demonstrating the thermal stabilities of PNAs of varying lengths bound to an $A_{10}/T_{10}$ double-stranded DNA target;

Figure 8 is a schematic model of C⁺G-G and T A-T triplets with Hoogsteen and Watson-Crick hydrogen bonds;

Figure 9 is a schematic model of the binding of PNA $(T_2CT_2CT_4)$ to a double-stranded DNA target;

Figure 10 shows a footprinting experiment of the binding of PNA to a dimeric target;

Figure 11 shows a footprinting experiment of PNA binding to a monomeric target;

Figure 12 is a footprinting experiment showing the effect of pH and salt on the binding of a thymine/cytosine containing PNA to double stranded DNA;

Figure 13 is an electron micrograph of a PNA double strand DNA complex;

Figure 14 is a histogram representation of the results obtained in the experiments of Figure 13;

Figure 15 is a one dimensional gel electrophoresis experiment showing the DNA unwinding upon PNA binding;

Figure 16 is a two dimensional gel electrophoresis experiment showing the DNA unwinding upon PNA binding;

Figure 17a shows the salt dependent binding of PNA in a gel electrophoresis experiment;

Figure 17b shows the salt resistance of the PNA-DNA complex in a gel electrophoresis experiment;

Figure 18 is an autoradiogram showing sequence selective transcription inhibition by PNA;

Figure 19 is an autoradiogram showing sequence specific inhibition of Taq DNA polymerase by PNA;

Figure 20 is an autoradiogram showing the concentration dependence of transcription inhibition by PNA $T_{10}$;

Figure 21 is a Figure 20 but using PNA $T_4CT_2CT_2$;

Figure 22 is a quantitative representation of the results of Figures 20 & 21;

Figure 23 is an autoradiogram showing the results of transcription arrest using various PNA and RNA polymerases;

Figure 24 is a continuation of Figure 23;

Figure 25 is an autoradiogram showing the specific effect

5  on transcription of PNA bound to the template versus the non-template strand;

Figure 26 is an autoradiogram showing transcription arrest by short ($T_6$&$T_8$) PNAs;

Figure 27 is a schematic drawing of PNA-DNA transcription

10  initiation complexes;

Figure 28 is an autoradiogram showing transcription initiation by PNA;

Figure 29 is an autoradiogram showing the efficient competition of a "PNA promoter" versus the strong lacUV5

15  promoter;

Figure 30 illustrates schematic structures of PNA/DNA complexes and sequences of the PNA targets;

Figure 31 is a plasmid map;

Figure 32 shows a PAGE autoradiography illustrating site

20  specific termination of in vitro transcription;

Figure 33 shows a PAGE autoradiography the effect of target site orientation on PNA mediated termination of in vitro transcription;

Figure 34 shows a PAGE autoradiography illustrating the

25  effect of salt concentration on PNA activity;

Figure 35 shows a PAGE autoradiography illustrating the effect of pH on PNA activity;

Figure 36 are charts plotting $K_D$ vs ph and $K_D$ vs ionic strength on in vitro binding of PNA to complement;

30  Figure 37 is a chart illustrating melting temperature for hybrid duplex binding;

Figure 38 is a chart illustrating melting temperature for hybrid duplex binding;

Figure 39 shows a PAGE autoradiography illustrating

35  titration by gel-shift binding of PNA to a end labeled oligonucleotide;

Figure 40 is a circular dichroism spectra;

- 17 -

Figure 41 is a further circular dichroism spectra;
Figure 42 is a chemical schematic; and
Figure 43 is a further chemical schematic.

## DETAILED DESCRIPTION OF THE INVENTION

5      For the purposes of this invention the term "purine rich"
shall mean a preponderance (i.e., greater than 50%) of purine
bases, preferably greater than about 65% and most preferrably
greater than about 90%. Further, the term "modify" shall mean
to change properties by addition to, substraction from,
10   cleavage of or otherwise, such that which results is
intrinsically different from that which is modified. The term
"modulate" shall mean to change (i.e., increase or decrease)
the magnitude of a property.

In accordance with the present inventions, the ligand L
15   is primarily a naturally occurring nucleobase attached at the
position found in nature, i.e., position 9 for purines
including adenine, guanine or inosine, and position 1 for
pyrimidines including thymine, uridine, or cytosine.
Alternatively, L may be a non-naturally occurring nucleobase
20   (nucleobase analog), another base-binding moiety, an aromatic
moiety, $(C_1-C_4)$alkanoyl, hydroxy or even hydrogen. It will be
understood that the term nucleobase includes nucleobases
bearing removable protecting groups. Some typical nucleobase
ligands and illustrative synthetic ligands are shown in Figure
25   2 of WO 92/20702. Two particular nucleobase ligands, 5-
propynylthymine and 3-deazauracil, have been shown to increase
binding affinity of an oligonucleotide to a target nucleotide
(see Froehler, B.C. et. al., Tetrahedron Letters, 1992 33:5307-
5310 and United States patent 5,134,066). Other like analogues
30   are shown by Sagi, J. et. al., Tetrahedron Letters, 1993
34:2191-2194. Incorporation of these nucleobase can be
effected to increase the binding affinity of the PNA product
with a nucleic acid target. Further useful non-naturally
occurring nucleobases include 6-thioguanine, i.e., purine-
35   6(1H)-thione, and pyrazolo[4,3-d]pyrimidines especially useful
as triplexing bases (see PCT application PCT/US92/04795).

Furthermore, L can be a DNA intercalator, a reporter ligand such as, for example, a fluorophor, a radio label, a spin label, hapten, or a protein-recognizing ligand such as biotin. In monomer synthons, L may be blocked with protecting

5  groups, as illustrated in Figure 4 of WO 92/20702.

Linker A can be a wide variety of groups such as $-CR^1R^2CO-$, $-CR^1R^2CS-$, $-CR^1R^2CSe-$, $-CR^1R^2CNHR^3-$, $-CR^1R^2C=CH_2-$ and $-CR^1R^2C=C(CH_3)_2-$, where $R^1$, $R^2$ and $R^3$ are as defined above. Preferably, A is methylenecarbonyl ($-CH_2CO-$), amido ($-CONR^3-$),

10  or ureido ($-NR^3CONR^3-$). Also, A can be a longer chain moiety such as propanoyl, butanoyl or pentanoyl, or corresponding derivative, wherein O is replaced by another value of X or the chain is substituted with $R^1R^2$ or is heterogenous, containing Y. Further, A can be a $(C_2-C_6)$alkylene chain, a $(C_2-C_6)$alkylene

15  chain substituted with $R^1R^2$ or can be heterogenous, containing Y. In certain cases, A can just be a single bond.

In one preferred form of the invention, B is a nitrogen atom, thereby presenting the possibility of an achiral backbone. B can also be $R^3N^+$, where $R^3$ is as defined above, or

20  CH.

In the preferred form of the invention, C is $-CR^6R^7-$, but can also be a two carbon unit, i.e., $-CHR^6CHR^7-$ or $-CR^6R^7CH_2-$, where $R^6$ and $R^7$ are as defined above. $R^6$ and $R^7$ also can be heteroaryl groups such as, for example, pyrrolyl, furyl,

25  thienyl, imidazolyl, pyridyl, pyrimidinyl, indolyl, or can be taken together to complete an alicyclic system such as, for example, 1,2-cyclobutanediyl, 1,2-cyclopentanediyl or 1,2-cyclohexanediyl.

In the preferred form of the invention, E in the monomer

30  synthon is COOH or an activated derivative thereof, and G in the oligomer is $-CONR^3-$ (amide). As defined above, E may also be CSOH, SOOH, $SO_2OH$ or an activated derivative thereof, whereby G in the oligomer becomes $-CSNR^3-$, $-SONR^3-$ and $-SO_2NR^3-$, (thioamide, sulfinamide or sulfonamide, respectively). The G

35  group can be in either orientation, e.g., for amide $-CONR_3-$ or $-R_3NCO-$. The activation may, for example, be achieved using an acid anhydride or an active ester derivative, wherein hydrogen

in the groups represented by E is replaced by a leaving group suited for generating the growing backbone.

The amino acids which form the backbone may be identical or different. We have found that those based on 2-aminoethyl-
5  glycine are especially well suited to the purpose of the invention.

In some cases it may be of interest to attach ligands at either terminus (Q, I) to modulate the binding characteristics of the PNAs. Representative ligands include DNA intercalators,
10  which improve dsDNA binding or basic groups, such as lysine or polylysine, which strengthen the binding of the PNA due to electrostatic interaction. To decrease electrostatic repulsion charged groups such as carboxyl and sulfo groups could be used. The design of the synthons further allows such other moieties
15  to be located on non-terminal positions. Oligonucleotides and/or oligonucleoside can be covalently bound to terminal positions Q or I to form chimeras containing PNA portions and oligonucleotide and/or oligonucleoside portions. Nucleosides and/or nucleotides (mono, di or tri-phosphates) also can be
20  attached to the terminal positions.

In a further aspect of the invention, the PNA oligomers are conjugated to low molecular weight effector ligands such as ligands having nuclease activity or alkylating activity or reporter ligands (fluorescent, spin labels, radioactive,
25  protein recognition ligands, for example, biotin or haptens). In a further aspect of the invention, the PNAs are conjugated to peptides or proteins, where the peptides have signaling activity and the proteins are, for example, enzymes, transcription factors or antibodies. Also, the PNAs can be
30  attached to water-soluble or water-insoluble polymers. In another aspect of the invention, the PNAs are conjugated to oligonucleotides or carbohydrates. When warranted, a PNA oligomer can be synthesized onto some moiety (e.g., a peptide chain, reporter, intercalator or other type of ligand-
35  containing group) attached to a solid support.

- 20 -

Such conjugates can be used for gene modulation (e.g., gene targeted drugs), for diagnostics, for biotechnology, and for scientific purposes.

As a further aspect of the invention, PNAs can be used
5 to target RNA and ssDNA to produce both antisense-type gene regulating moieties and hybridization probes for the identification and purification of nucleic acids. Furthermore, the PNAs can be modified in such a way that they form triple helices with dsDNA. Reagents that bind sequence-specifically
10 to dsDNA have applications as gene targeted drugs. These are foreseen as extremely useful drugs for treating diseases like cancer, AIDS and other virus infections, and may also prove effective for treatment of some genetic diseases. Furthermore, these reagents may be used for research and in diagnostics for
15 detection and isolation of specific nucleic acids.

Triple helix formation, wherein an oligonucleotide is triplexed to a dsDNA, is believed to be the only means known in the art for sequence-specific recognition of dsDNA. However, triple helix formation is largely limited to recognition of
20 homopurine-homopyrimidine sequences. Triplexing with strand displacement using PNAs of this invention is superior to oligonucleotide-dsDNA triple helix recognition in that it may allow for recognition of any sequence by use of the four natural bases.

25 Gene targeted drugs are designed with a nucleobase sequence (containing 10-20 units) complementary to the regulatory region (the promoter) of the target gene. Upon administration, the drug binds to the promoter and blocks access thereto by RNA polymerase. Consequently, no mRNA, and
30 thus no gene product (protein), is produced. If the target is within a vital gene for a virus, no viable virus particles will be produced. Alternatively, if the target is downstream from the promoter, RNA polymerase will terminate at this position, thus forming a truncated mRNA/protein which is nonfunctional.
35 Sequence-specific recognition of ssDNA by base complementary hybridization can be exploited to target specific genes and viruses. In this case, the target sequence is contained in

the mRNA such that binding of the drug to the target hinders
the action of ribosomes and, consequently, translation of the
mRNA into protein.  The peptide nucleic acids of the invention
are superior to prior reagents in that they have significantly
5    higher affinity for complementary ssDNA.  Also while a charged
species such as a lysine moiety can be added, PNAs possess no
charge, they are water soluble (which should facilitate
cellular uptake), and they contain amides of non-biological
amino acids (which should make them biostable and resistant to
10   enzymatic degradation by, for example, proteases).  Further,
they can triplex with the mRNA.

The PNA compounds used in this invention can be
synthesized by the methodologies disclosed in WO 92/20702,
WO/92/20703 and the foregoing United States patent application
15   bearing attorney docket ISIS-1017.  Monomer synthons according
to those disclosures are coupled using standard protocols to
give the desired PNA oligomeric sequences.  The synthesis of
additional synthetic monomers are given by example 33 through
45 of this specification.

20       Binding of PNA compounds to double-stranded DNA
accompanied by strand displacement is shown in illustrative
examples in this specification by both enzymatic and chemical
probing.  Our illustrative examples show that the PNA compounds
bind via Watson-Crick binding to their complementary strand and
25   extrudes the other strand in a virtually single-stranded
conformation.

Using the PNA compound to locally unwind the DNA duplex
and effect strand displacement can render the displaced DNA
strand sensitive to cleavage (e.g., cleavage by $S_1$ nuclease).
30   This gives or induces high yield double strand DNA breaks at
the site of PNA binding.  PNA-directed and PNA-provoked double
strand DNA cleavage by $S_1$ nuclease is especially efficient when
two adjacent PNA sites are targeted, yielding quanti-tative
conversion of the target site into double-strand breaks.

35       As an illustrative example of use as an artificial
restriction enzyme, the targets are cloned within a pUC19
polylinker and the plasmids are linearized with the Cfr10I

- 22 -

restriction enzyme.   In the linear DNA, the resulting
polylinker region is in the middle (1.33 kb from one end and
1.36 kb from the other).   A pT10 plasmid carrying a known
sequence is then inserted in the unique BamHl site in the
5    polylinker.   This is complexed with a complementary PNA
compound and subjected to treatment with $S_1$ nuclease.   Upon
nuclease treatment, a significant fraction of DNA is cut, as
shown by gel electrophoresis of the resulting fragments.  Cross
reactivity experiments show that the targeting is sequence-
10   specific; only corresponding PNAs mediated cutting of the
targets.   The yield of digested molecules is high and one can
reach the quantitative digestion with increasing exposure to
nuclease.

        The results of such treatment demonstrate a novel method
15   for sequence-selective cleavage of double-stranded DNA.  If two
closely positioned homo-pyrimidine stretches can be identified
(as for example 7-10 bases, preferably on opposite strands),
this can be done by synthesizing pairs of PNAs complementary to
two parts of this DNA sequence and separated by several base
20   pairs.   If this strand displacement binding mode is extended to
include PNAs recognizing DNA sequences containing thymine and
cytosine, this strategy allows targeting and specific cleavage
of any desired sequence of 10-20 base pairs.   In effecting such
cleavage, the high stability of PNA-DNA complexes consisting of
25   two PNA strands (one "Watson-Crick" and one "Hoogsteen") and
one DNA strand must be considered, as must the sequence of the
"Hoogsteen-like" strand.

        We have found that complexes between homopyrimidine PNA
and complementary oligonucleotides show 2:1 stoichiometry.
30   Further we have found that the thermal stability of triplexes
between   cytosine-containing   homopyrimidine   PNA   and
complementary oligonucleotides is strongly pH dependent,
indicating the involvement of C⁺G-C Hoogsteen hydrogen bonding
in $(PNA)_2$/DNA triplexes.   Further, while we do not wish to be
35   bound by any particular theory, we believe that the strand
displacement binding of homopyrimidine PNA to dsDNA is
dependent on the formation of such $(PNA)_2$/DNA triplexes with

- 23 -

the complementary strand of the dsDNA target. It is our present belief that formation of the strand displacement complex proceeds via a DNA/PNA Watson-Crick duplex which is subsequently stabilized by a Hoogsteen bound second strand
5 having C$^+$-G Hoogsteen hydrogen bonds.

In further studies of triplexing and strand displacing, we utilized a cytosine containing homopyrimidine PNA ($T_2CT_2CT_4$-LysNH$_2$; SEQ ID NO:1) and studied its binding to a dimeric target. Stability of the complex was probed by both
10 dimethylsulfate (DMS) probing and KMnO$_4$ probing. In these studies, we found that triple strand formation involves Hoogsteen hydrogen bonding. This was evidenced by our showing that the N7 atoms of guanines of the target participate in base pairing and, thus, are protected from reaction with DMS upon
15 triplexing binding of the PNA. A model of this triplexing binding is shown in Figures 8 and 9.

We have also used electron microscopy to confirm strand displacement. In these examples, a PNA-DNA complex was formed with a target. Full occupancy of the target resulted in a
20 strand displacement loop of 90-100 bases, as detected by electron microscopy. The DNA molecules carry an open region in the form of an "eye". In all cases, one of the two strands in the open region is thicker than the other one and has the same thickness as the normal DNA duplex. The thicker strand
25 corresponds to the strand covered by PNA while the thinner strand corresponds to the displaced strand. In control experiments carried out without added PNA, the displaced structures are not observed.

Discrimination of PNA compounds in binding to targets was
30 also observed utilizing labeled DNA fragments that were
challenged with various PNA compounds. We show that PNA compounds bind with high preference to the complementary target. Binding decreases with increasing mismatches in the targets. Binding to a target containing one mismatch was
35 weaker than for a fully complementary target, while no binding was observed with a two-mismatch target.

- 24 -

We have further found that PNA compounds truly mimic DNA in terms of base pair specific hybridization to complementary strands of oligonucleotides as measured by thermal stability of complexes. We have shown that the formation of PNA/DNA
5  duplexes exhibits a decrease in entropy almost identical to that for the formation of DNA/DNA duplexes and that like the formation of a DNA/DNA it is strongly enthalpy driven. From this we deduce that single stranded PNA must have the same degree of base stacking as single stranded DNA and thus appears
10  to be highly structured. Further we have found that the rate of hybridization of PNA/RNA duplexes is at least as fast as that for 2'-O-methyl RNA/RNA or DNA/DNA duplex formation again supporting our findings that single stranded PNA is at least as prestructured for duplex formation as is DNA or RNA.
15      We have further found that in contrast to DNA or RNA, PNA compounds may bind to complementary DNA or RNA in one of two orientations, a parallel orientation or an anti-parallel orientation. We have further found that the PNA compounds prefer the anti-parallel orientation (wherein the amino end of
20  the PNA strand is complementary to the 3'-end of the nucleic acid) over the parallel orientation ((wherein the amino end of the PNA strand is complementary to the 5'-end of the nucleic acid).

In our studies we have shown that homo-pyrimidine PNA
25  binds strongly to ssDNA, dsDNA or RNA with 2:1 stoichiometry (triplexing) and effects stable strand displacement complexes with dsDNA. Further, purine-pyrimidine PNA binds to Watson-Crick complementary oligonucleotides, either RNA or DNA, with 1:1 stoichiometry. Utilizing these findings, one embodiment of
30  this invention is directed to use of PNA compounds as effective inhibitors of the elongation activity of RNA and DNA polymerases. As contrasted to oligonucleotides, this allows use of PNA compounds to target genes not only at promoter or regulatory regions of the genome but to other regions as well.
35  Furthermore, only decamer targets are required for efficient binding of the PNA to dsDNA. Consequently, it is easier to

identify suitable PNA targets as compared to oligonucleotide triple helix targets.

We have further found that PNA compounds have certain effects on transcription in both prokaryotic and eukaryotic systems. Transcription by RNA polymerases $T_3$ or $T_7$ of PNA $T_{10}$-LysNH$_2$ (SEQ ID NO:2) and $T_2CT_2CT_4$-LysNH$_2$ (SEQ ID NO:1) bound to dsDNA targets $A_{10}$ (SEQ ID NO:3), $A_5GA_4$ (SEQ ID NO:4) and $A_2GA_2GA_4$ (SEQ ID NO:5) positioned downstream from $T_3$ or $T_7$ promoters in pBluescriptKS$^+$ plasmids were studied. As shown in our illustrative examples below, transcription elongation is arrested at the site of PNA binding to the template strand, whereas only a marginal effect is observed at the site of PNA binding to the non-template strand. With PNA $T_{10}$-LysNH$_2$ (SEQ ID NO:2), transcription arrest occurs at the first base of the PNA binding site, while arrest with PNA $T_5CT_4$-LysNH$_2$ (SEQ ID NO:6) takes place 2-3 bases inside the PNA binding site. In the case of PNA $T_2CT_2CT_4$-LysNH$_2$ (SEQ ID NO:1), arrest is less efficient and occurs at the last 1-3 bases of the binding site.

Transcription arrest has also been shown for PNAs $T_6$-LysNH$_2$ (SEQ ID NO:7) and $T_8$-LysNH$_2$, although at lower efficiency compared to longer PNA compounds. These results show that efficient transcription elongation arrest can be obtained by PNA targeting of the transcribed strand, and that "read through" by the polymerase takes place in a sequence dependent manner.

Transcription arrest in eukaryotic systems has further been shown for PNA compounds. For these studies plasmids were constructed with the CMV IE 1 promoter driving the transcription of viral DNA sequences. The viral DNA sequences containing homopurine target sites for the PNAs were cloned downstream of the CMV IE 1 promoter. The plasmids were incubated with PNA oligomers under various conditions then added to eukaryotic nuclear extracts to initiate in vitro transcription. Transcripts were specifically truncated at the site of PNA binding. This effect was dependent upon hybridization of the PNA to the template strand, i.e., the inhibition of transcription was sequence specific. Binding

- 26 -

affinity was increased by either lowering the ionic strength or pH of the binding buffer. These results are consistent for PNA binding in which homopyrimidine PNA's invade duplex DNA binding the complementary DNA strand, while displacing the non-
5   complementary strand. A second PNA then interacts with the PNA/DNA duplex forming a stable triple stranded structure which is capable of blocking RNA polymerase II and terminating transcription thus specifically inhibiting eukaryotic gene expression. This use of PNA compounds to inhibit eukaryotic
10  gene expression renders the PNA compounds useful for antigene therapeutics.

     In a further embodiment of this invention, we have found that PNA compounds also can be used as sequence specific gene activators and synthetic transcription factors. Transcription
15  initiation by RNA polymerase involves the sequence specific recognition of the double-stranded DNA promoter either by the polymerase itself or by auxiliary transcription factors. Subsequently a transcription initiation open complex is formed in which about 12 base pairs of the DNA helix are melted. This
20  exposes the bases of the template strand for base pairing with the RNA strand being synthesized. It has been shown that an E. coli phage T₇ RNA polymerase can utilize synthetic "RNA/DNA bubble duplex" complexes containing an RNA/DNA duplex and a single-stranded DNA D-loop for transcription elongation. We
25  have found that homopyrimidine PNAs also form D-loop structures when binding to complimentary double-stranded DNA by strand displacement. We believe that these structures behave like RNA/DNA open complex structures and are recognized by RNA polymerase.
30       In the illustrative examples below, we show that E. Coli RNA polymerase does indeed bind to PNA/dsDNA strand displacement complexes and initiates RNA transcription therefrom. The results of these examples further suggest that a single-stranded DNA loop is a major structural determinant
35  for RNA polymerase upon transcription initiation and elongation. The results of these examples also have

implications for elucidating the mechanism of action of RNA polymerase.

In further embodiments of the invention chimera strands are formed between PNAs and either RNA or DNA. Such chimeric
5  strands can then in the same manner as the "homo" PNA strands described above. Such chimeric structures thus will be used, as described above, for binding, duplexing, triplexing and the like. In one particularly preferred use, either a PNA strand or a PNA containing chimera strand will be used to bind to or
10  otherwise modulate proteins in cells. Such proteins will include transcription factors and other regulatory proteins.

The chimeric structure between PNAs and DNA or RNA are used in place of or in addition to a normal PNA strand to effect duplexing, triplexing, nucleic acid binding or protein
15  binding. The RNA or DNA nucleic acid portion of such chimeric structures include nucleic acid connected via phosphodiester, phosphorothioate, phosphorodithioate, alkyl phosphonate, hydrogen phosphonate, phosphotriester, phosphoramidite and other like phosphorus linkages. They further can include other
20  substitutions such as substitution at the 2' position of a ribose sugar. Particularly preferred are 2'-deoxy-2'-fluoro since they increase affinity of the nucleic acid portion of the chimera to other nucleic acids and 2'-O-alkyl, particularly 2'-O-propyl, 2'-O-allyl and the like since they confer nuclease
25  resistance to the nucleic acid strand.

The following examples are given to illustrate the invention. These examples are given for illustrative purposes and are not meant to be limiting.

**Example 1**
30  **General Method for the Synthesis of PNA Oligomers**

PNA oligomeric compounds were prepared generally in accordance with the methods disclosed by WO 92/20702, WO 92/20703 and the foregoing United States patent application bearing attorney docket ISIS-1017. Other PNA monomers are
35  prepared as per Examples 34-46 below. Briefly, benzyhydrylamine resin (initially loaded 0.28 mmol/gm with Boc-L-Lys(2-chlorobenzyloxycarbonyl)) was swollen in DMF and an

excess of a monomer to be coupled was added, followed by dicyclohexylcarbodiimide (0.15M in 50% DMF in dichloromethane). The Boc deprotection was accomplished by trifluoroacetic acid treatment. The progress of the coupling reactions was
5   monitored by qualitative or quantitative ninhydrin analysis. The PNA was released from the resin using anhydrous HF or trifluoromethyl sulfonic acid under standard conditions. The products were purified using HPLC with acetonitrile-water (0.1%TFA) gradient and structure confirmed by fast atom
10  bombardment mass spectrometry.

Representative sequences synthesized by these methods include the following as well as other sequences noted in the various examples:

H-$T_{10}$LysNH$_2$ (SEQ ID NO:2)
15  H-$T_4$CT$_5$LysNH$_2$ (SEQ ID NO:8)
H-$T_2$CT$_2$CT$_4$LysNH$_2$ (SEQ ID NO:1)
H-$T_4$CT$_2$CT$_2$LysNH$_2$ (SEQ ID NO:9)
H-TGTACGTCACAACTA-NH$_2$ (SEQ ID NO:10)
H-CCTTCCCTT-NH$_2$ (SEQ ID NO:11)
20  H-TTCCCTTCC-NH$_2$ (SEQ ID NO:12)
H-TAGTTATCTCTATCT-NH$_2$ (SEQ ID NO:13)
H-TGTACGTCACAACTA-NH$_2$ (SEQ ID NO:14)
H-GCACAGCC-LYS-NH$_2$ (SEQ ID NO:15)
H-TTTTCTTTT-NH$_2$ (SEQ ID NO:16)
25  H-TTTTTTTTTCCCCCCC-NH$_2$ (SEQ ID NO:17)
H-CCCCCCCTTTTTTTTT-NH$_2$ (SEQ ID NO:18)
H-CCTCCTTCCC-NH$_2$ (SEQ ID NO:19)
H-TTCTCTCTCT-NH$_2$ (SEQ ID NO:20)
H-TTTTTCTCTCTCTCT-NH$_2$ (SEQ ID NO:21)
30  H-CCCCCACCACTTCCCCTCTC-(Lys)$_9$NH$_2$ (SEQ ID NO:22)
H-CTTATATTCCGTCATCGCTC-Lys-NH$_2$ (SEQ ID NO:23)
H-CTGTCTCCATCCTCTTCACT-NH$_2$ (SEQ ID NO:24)
H-TATTCCGTCATCGCTCCTCA-Lys-NH$_2$ (SEQ ID NO:25)
H-CCCCCACCACTTCCCCTCTC-NH$_2$ (SEQ ID NO:26)
35  H-CTGCTGCCTCTGTCTCAGGT-LysNH$_2$ (SEQ ID NO:27)
H-$T_4$-(ß-alanine)C-$T_5$-LysNH$_2$ (SEQ ID NO:28)
H-$T_4$-($\beta$-alanine)T-$T_5$-LysNH$_2$ (SEQ ID NO:29)

- 29 -

The PNA's are written from amino to the carboxyterminal. LysNH$_2$ designates a lysine amide is attached to the PNA and NH$_2$ indicates a free c-terminal carboxamide without lysine.

**EXAMPLE 2**

5  **Site-specific S$_1$ nuclease digestion of the pTlO plasmid linearized with CfrlOI restriction enzyme and complexed with PNA (a double-stranded target)**

PNA H-T$_{10}$-LysNH$_2$ (SEQ ID NO:2) was synthesized as described in Example 1.  A pTlO plasmid was prepared from the
10  pUC19 plasmid by inserting the dA$_{10}$/dT10 (SEQ ID NO:3) sequence into the BamHl site of the polylinker as per the method of Egholm, M et al., *J.A.C.S.* 1992 *114*:1895-1897.  The pTlO plasmid was linearized with CfrlOI restriction enzyme in the unique site.  To form complex with the PNA, about 0.1 µg of the
15  linearized plasmid was incubated with 2 o.u./ml of PNA in 3 µl of the TE buffer (10 mM Tris-HCI; 1 mM EDTA, pH 7.4) at 37 °C. To perform the S$_1$ nuclease reaction, 10 µl of Na-Acetate buffer (33 mM NaAc; 50 mM NaCl; 10 mM ZnSO$_4$; 0.5% of glycerol, pH 4.6) and 145 units of S$_1$ nuclease (Sigma) were added and incubated
20  for various time periods at room temperature. The reaction was terminated by adding 1 µl of 0.5 M EDTA and cooling to -20 °C. Electrophoresis was performed in 1% agarose gel in the TBE buffer with subsequent staining with ethidium bromide.  The results of this example are shown in Figure 1 wherein the size
25  of the markers are shown as are the incubation time periods. In Figure 8 Lane 1: control - preincubation of DNA with PNA was performed under conditions unfavorable for complex formation (200 mM of NaCl). Lane 2: the reference band obtained by digestion with BamHI restriction enzyme. Lane 3: reference
30  bands obtained by the digestion with BglI restriction enzyme. Lanes 4,5: the results of S$_1$ nuclease digestion for different times of DNA-PNA complex.

As is shown in this example, targets were cloned within a pUC19 polylinker.  The plasmids were linearized with the
35  CfrlOI restriction enzyme.  In the linear DNA, the obtained polylinker region is in the middle (1.33 kb from one end and 1.36 kb from the other).  Utilizing a direct approach, a pTl0

- 30 -

plasmid carrying the sequence $dA_{10}dT_{10}$ (SEQ ID NO:31) was inserted in the unique BamH1 site in the polylinker. This was complexed with PNA $H-T_{10}-LysNH_2$ (SEQ ID NO:2) and subjected to treatment with $S_1$ nuclease. After treatment with 145 units of

5    $S_1$ nuclease at room temperature a significant fraction of DNA was cut. The mobility by gel electrophoresis of the resulting fragments was very close to the mobility of fragments obtained by cutting with restriction enzyme BamH1. Similar results were obtained for $H-T_5CT_4-LysNH_2$ (SEQ ID NO:6) and $H-T_2CT_2CT-LysNH_2$

10   (SEQ ID NO:51) PNAs and corresponding plasmids carrying the $d(A_5GA_4)/d(T_4CT_5)$ (SEQ ID NO:4) and $d(A_4GA_2GA_2)/d(T_2CT_2CT_4)$ (SEQ ID NO:44) inserts. Cross reactivity experiments showed that the targeting was sequence specific: only corresponding PNAs mediated cutting of the targets. It was noted that two very

15   week bands are seen in lanes 4,5, which correspond to lengths 2.32 and 0.98 kb. These bands are due to weak binding of PNA-TIO to the intrinsic pUC19 site d(TTGT3)/d(A3CAT), which is 0.37 kb apart from the Cfr10I restriction site. This is consistent with the results of Example 3 above that shows that

20   the introduction of mismatches dramatically decreases the affinity of PNA to DNA.

From this example it is believed that the enzyme first digests the displaced strand, then to some extent enlarges the gap after which the opposite strand becomes a substrate for the

25   enzyme. As a result, the double-stranded break is created. The yield of digested molecules is high and one can reach the quantitative digestion with increasing exposure to $S_1$ nuclease (data not shown). However, while the two fragments are poorly resolved after digestion with BamHI restriction enzyme, the

30   PNA-mediated digestion with the $S_1$ nuclease leads to the clear-cut doublet. While not wishing to be bound by theory, we believe this probably reflects a widening of the gap by the $S_1$ nuclease as well as digestion from the ends of the cfr10I site. With longer treatment the downward shift of the doublet becomes

35   noticeable (see lane 5). Thus, quantitative digestion is clearly accompanied by a truncation of the fragments (by about hundred base pairs). To obtain further precision of cutting

- 31 -

comparable with that exhibited by natural restriction enzymes the original approach is modified as per Example 4 below.

**Alternate site-specific S₁ nuclease digestion of plasmid with PNA)**

5       As an alternative target, the insert:

           5'-A₅GA₄GTCGACA₅GA₄-3' (SEQ ID NO:30)
           3'-T₅CT₄CAGCTGT₅CT₄-5'

was cloned into the SalI site of the pUC19 plasmid.   This plasmid, which contained two binding sites for PNA H-T₅CT₄-

10  LysNH₂  (SEQ  ID  NO:6)  separated  by  six  base  pairs,  was designated pT9CT9C.  Strand displacement in the two T₅CT₄ sites led  to  opening  of  the  entire  region,  including  the  sequence GTCGAC/GACGAC,  providing  a  substrate  for  the  S₁  nuclease  in both strands.  As control, a pT9C-5 plasmid was used, which

15  carried  only  one  T₅CT₄  insert  cloned  into  the  SalI  site  of  the pUC19 plasmid.


EXAMPLE 3

Site-specific  S₁  nuclease  digestion  of  the  pT9C  plasmid

20  **linearized by CfrlOI restriction enzyme and complexed with PNA H-T₅CT₄-LysNH₂**

           This example illustrates the use of one molecule of PNA hybridized to a target DNA to define a restriction site.

           PNA 1,2 and 3 were synthesized as Example 1 and the

25  results  are  shown  in  Figure  2.    A  pT9C  plasmid  carried  the insert:

           5 '-A₅GA₄GTCGACA₅GA₄-3' (SEQ ID NO:30)
           3 '-T₅CT₄CAGCTGT₅CT₄-5',

cloned in the SalI site of the pUC19 polylinker.  The pT9C-5

30  plasmid  carried  the  single  insert  A₅GA₄/CT₅  cloned  in  the  same site.    The  PNA-DNA  complexes  were  prepared  as  described  in Example 2 with the only difference being the duration of the incubation was 2 hours.  Digestion by 30 units of the S₁ nuclease  was  performed  in  the  same  buffer  as  described  in

35  Example 2 with two exceptions, lanes 2 and 4.  In lane 2, 15 units of the enzyme were used, whereas in lane 4 200 mM NaCl and 1 mM of ZnSO₄ were added to the buffer.  Lane 1: the

- 32 -

reference band obtained by digestion with the BamHI restriction
enzyme of the pUC19 plasmid linearized by the Cfr10I
restriction enzyme. Lanes 2-4 site-specific digestion by the
S$_1$ nuclease of the linearized pT9C plasmid complexed with PNA
5    H-T$_5$CT$_4$-LysNH$_2$ (SEQ ID NO:6). Lanes 5-7: various controls.
Lane 8: the same experiment as in lane 3, but using PNA T$_{10}$ (SEQ
ID NO:2) instead of T$_4$CT$_5$ (SEQ ID NO:8). Lane 9: the same
experiment as in lane 3, but using PNA T$_2$CT$_2$CT$_4$ (SEQ ID NO:1)
instead of T$_4$CT$_5$ (SEQ ID NO:8). Lane 10: the same experiment
10   as in lane 3, but using the pT9C-5 plasmid instead of the pT9C.
In the header PNA T$_5$CT$_4$ (SEQ ID NO:6) is labelled as 1, T$_{10}$ (SEQ
ID NO:2) as 2 and T$_2$CT$_2$CT$_4$ (SEQ ID NO:1) as 3.

     This example show that subjecting the pT9CT9C plasmid,
linearized by the Cfr10I restriction enzyme and complexed with
15   PNA H-T$_5$CT$_4$-LysNH$_2$ (SEQ ID NO:6), to 30 units of the S$_1$ nuclease
results in full conversion of the full-length DNA molecules
into half-length fragments. The bands observed by gel
electrophoresis have the same position and width as the band
created by the BamHI restriction enzyme. Sequence specificity
20   of this artificial "restriction enzyme" was confirmed by the
very weak digestion of the  pT9C plasmid in case of complexing
with PNA H-T$_{10}$-LysNH$_2$ (SEQ ID NO:2) and H-T$_2$CT$_2$CT$_4$-LysNH$_2$ (SEQ ID
NO:1) under conditions which result in quantitative cutting of
the plasmid in the presence of PNA H-T$_5$CT$_4$-LysNH2 (SEQ ID NO:6)
25   (see lanes 8,9). Moreover, under the much milder S$_1$ nuclease
treatment necessary to generate the data for the pT9CT9C, the
yield of double-stranded breaks in the pT9C-5 plasmid was
extremely low (lane 10).
**EXAMPLE 4**
30   **Site specific S$_1$ nuclease cleavage of the plasmids pT9C-5,
pT9CT9C and pT9CA9GKS (linearized with ScaI) targeted by PNA
T$_4$CT$_5$-LysNH$_2$**

     This example illustrates the use of two molecules of PNA
bound to opposite strands of a DNA target. A further plasmid
35   pT9CA9GKS    was    produced    by    cloning    the    insert
GTCGACA$_5$GA$_4$GTCGACT$_4$CT$_5$GTCGAC (SEQ ID NO:32) into pUC19 at the
Sal I site and linearizing with Sca I restriction enzyme. The

linearized plasmid has two hybridization sites for PNA H-T$_4$CT$_5$-Lys NH$_2$ (SEQ ID NO:8) on opposite strands spaced by six base pairs. The protocol of Example 3 was use except the samples were treated using 1 U/$\mu$l of S$_1$ and 15 min incubation at 37 °C.

5 The results are shown in Figure 3 Lanes 1-4: pT9C-5; lanes 5-8: pT9CT9C; lanes 9-12: pT9CA9GKS. Lanes 1, 5 & 9: no PNA; lanes 2, 6 & 10: 50 $\mu$M; lanes 3, 7, 11: 500 uM; lanes 4, 8, 12: 5 mM.

**Example 5**

**Use of PNA/nuclease to cut selectively large DNA molecule**

10       The standard strain lambda cI indl ts857 Sam7 (New England Biolab) was used. The PNA double target was T$_7$N$_6$T$_7$ (SEQ ID NO:33) which upon cleavage should give rise to a 6.1 kb fragment. The PNA was a PNA (T$_7$)$_2$ and the conditions were as follows: 0.1 $\mu$g DNA and 1-2 OD units of PNA were incubated in

15 5 $\mu$l TE buffer for 10 min at 37 °C. 5 $\mu$l buffer (33 mM NaAc, 30 mM NaCl, 10 $\mu$M ZnSO$_4$, pH 4.5) and 20 U mung bean nuclease were added and the mixture was incubated for 5 min at room temperature. Upon analysis by electrophoresis in 0.5% agarose in TBE buffer, fragments of sizes 6 kb and 42 kb were observed

20 using lambda x _Hind_ III or _Sal_ I as size markers. Without additions of PNA no bands were seen.

       This example demonstrates the ability of PNA/nuclease to cut selectively a large DNA molecule (by comparison to the much smaller plasmids used in Examples 1 to 3).

25 **EXAMPLE 6**

**Binding of PNAs-T$_{10}$/T$_9$C/T$_8$C$_2$ to double-stranded DNA targets A$_{10}$/A$_9$G/A$_8$G$_2$**

       A mixture of 200 cps $^{32}$P-labeled _EcoRI-PvuII_ fragment (the large fragment labeled at the 3'-end of the _EcoRI_ site) of

30 the indicated plasmid, 0.5 $\mu$g carrier calf thymus DNA, and 300 ng PNA in 100 $\mu$l buffer (200 mM NaCl, 50 mM Na-acetate, pH 4.5, 1 mM ZnSO$_4$) was incubated at 37°C for 120 min. A 50 unit portion of nuclease S$_1$ was added and incubated at 20°C for 5 min. The reaction was stopped by addition of 3 $\mu$l 0.5 M EDTA

35 and the DNA was precipitated by addition of 250 $\mu$l 2% potassium acetate in ethanol. The DNA was analyzed by electrophoresis in

10% polyacrylamide sequencing gels and the radiolabeled DNA bands visualized by autoradiography.

The target plasmids were prepared by cloning of the appropriate oligonucleotides into pUC19. Target $A_{10}$: oligo-
5    nucleotides GATCCA$_{10}$G (SEQ ID NO:34) & GATCCT$_{10}$G (SEQ ID NO:35) cloned into the BamHI site (plasmid designated pT10). Target $A_5GA_4$: oligonucleotides TCGACT$_4$CT$_5$G (SEQ ID NO:36) & TCGACA$_5$GA$_4$G (SEQ ID NO:37) cloned into the SalI site (plasmid pT9C). Target $A_2GA_2GA_4$: oligonucleotides CA$_2$GA$_2$GA$_4$CTGCA (SEQ ID NO:38)
10   and GT$_4$CT$_2$CT$_2$CTGCA (SEQ ID NO:39) into the PstI site (plasmid pT8C2). The results are shown in Figure 4. The positions of the targets in the gel are indicated by bars to the left. A/G is an A+G sequence ladder of target P10.
**EXAMPLE 7**
15   **Inhibition of restriction enzyme cleavage by PNA (Figure 4)**

A 2 $\mu$g portion of plasmid pT10 was mixed with the indicated amount of PNA-T$_{10}$ (SEQ ID NO:2) in 20 $\mu$l TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.4) and incubated at 37°C for 120 min. 2 $\mu$l 10 × buffer (10 mM Tris-HCl, pH 7.5, 10 mM, MgCl$_2$, 50
20   mM NaCl, 1 mM DTT). PvuII (2 units) and BamHI (2 units) were added and the incubation was continued for 60 min. The DNA was analyzed by gel electrophoresis in 5% polyacrylamide and the DNA was visualized by ethidium bromide staining. This gel is illustrated in Figure 5.
25   **EXAMPLE 8**
**Kinetics of PNA-T$_{10}$ - dsDNA strand displacement complex formation**

A mixture of 200 cps $^{32}$P-labeled EcoRI-PvuII fragment of pT10 (the large fragment labeled at the 3'-end of the EcoRI
30   site), 0.5 $\mu$g carrier calf thymus DNA, and 300 ng of PNA-T$_{10}$-LysNH$_2$ (SEQ ID NO:2) in 100 $\mu$l buffer (200 mM NaCl, 50 mM Na-acetate, pH 4.5, 1 mM ZnSO$_4$) were incubated at 37°C. At the times indicated, 50 U of S$_1$ nuclease was added to each of 7 samples and incubation was continued for 5 min at 20°C. The
35   DNA was then precipitated by addition of 250 $\mu$l 2% K-acetate in ethanol and analyzed by electrophoresis in a 10% polyacrylamide sequencing gel. The amount of strand displacement complex was

- 35 -

calculated from the intensity of the $S_1$-cleavage at the target
sequence, as measured by densitometric scanning of
autoradiographs. The results are illustrated in Figure .6.

**EXAMPLE 9**

5  **Stability of PNA-dsDNA complexes**

A mixture of 200 cps $^{32}$P-pT10 fragment, 0.5 $\mu$g calf
thymus DNA and 300 ng of the desired PNA (either $T_{10}$-LysNH$_2$ (SEQ
ID NO:2), $T_8$-LysNH$_2$ (SEQ ID NO:40) or $T_6$-LysNH$_2$ (SEQ ID NO:7))
was incubated in 100 $\mu$l 200 mM NaCl, 50 mM Na-acetate, pH 4.5,

10  1 mM ZnSO$_4$ for 60 min at 37°C. A 2 $\mu$g portion of
oligonucleotide GATCCA$_{10}$G (SEQ ID NO:34) was added and each
sample was heated for 10 min at the temperature indicated,
cooled in ice for 10 min and warmed to 20°C. A 50 U portion of
$S_1$ nuclease was added and the samples treated and analyzed and

15  the results quantified as is illustrated in Figure 7.


**EXAMPLE 10**

**Biological stability of PNA**

A mixture of PNA-T$_5$ (SEQ ID NO:41) (10 $\mu$g) and a control,

20  "normal" peptide (10 $\mu$g) in 40 $\mu$l 50 mM Tris-HCl, pH 7.4 was
treated with varying amounts of peptidase from porcine
intestinal mucosa or protease from Streptomyces caespitosus for
10 min at 37°C. The amount of PNA and peptide was determined
by HPLC analysis (reversed phase C-18 column: 0-60%

25  acetonitrile, 0.1% trifluoroacetic acid).

At peptidase/protease concentrations where complete
degradation of the peptide was observed (no HPLC peak) the PNA
was still intact.

**EXAMPLE 11**

30  **Inhibition of Gene Expression**

A preferred assay to test the ability of peptide nucleic
acids to inhibit expression of the E2 mRNA of papillomavirus is
based on the well-documented transactivation properties of E2.
Spalholtz, et al., J. Virol., 1987, 61, 2128-2137. A reporter

35  plasmid (E2RECAT) was constructed to contain the E2 responsive
element, which functions as an E2 dependent enhancer. E2RECAT
also contains the SV40 early promoter, an early polyadenylation

signal, and the chloramphenicol acetyl transferase gene (CAT).
Within the context of this plasmid, CAT expression is dependent
upon expression of E2. The dependence of CAT expression on the
presence of E2 has been tested by transfection of this plasmid
5  into C127 cells transformed by BPV-1, uninfected C127 cells and
C127 cells cotransfected with E2RECAT and an E2 expression
vector.

### A.    Inhibition of BPV-1 E2 Expression

BPV-1 transformed C127 cells are plated in 12 well
10 plates. Twenty four hours prior to transfection with E2RE1,
cells are pretreated by addition of antisense PNAs to the
growth medium at final concentrations of 5, 15 and 30 mM. The
next day cells are transfected with 10 $\mu$g of E2RE1CAT by
calcium phosphate precipitation. Ten micrograms of E2RE1CAT
15 and 10 $\mu$g of carrier DNA (PUC 19) are mixed with 62 $\mu$l of 2 M
$CaCl_2$ in a final volume of 250 $\mu$l of $H_2O$, followed by addition
of 250 $\mu$l of 2X HBSP (1.5 mM $Na_2PO_2$. 10 mM KCl, 280 mM NaCl, 12
mM glucose and 50 mM HEPES, pH 7.0) and incubated at room
temperature for 30 minutes. One hundred microliters of this
20 solution is added to each test well and allowed to incubate for
4 hours at 37 °C. After incubation, cells are glycerol shocked
for 1 minute at room temperature with 15% glycerol in 0.75 mM
$Na_2PO_2$, 5 mM KCl, 140 mM NaCl, 6 mM glucose and 25 mM HEPES, pH
7.0. After shocking, cells are washed 2 times with serum free
25 DMEM and fed with DMEM containing 10% fetal bovine serum and
antisense oligonucleotide at the original concentration. Forty
eight hours after transfection cells are harvested and assayed
for CAT activity.

For determination of CAT activity, cells are washed 2
30 times with phosphate buffered saline and collected by scraping.
Cells are resuspended in 100 $\mu$l of 250 mM Tris-HCl, pH 8.0 and
disrupted by freeze-thawing 3 times. Twenty four microliters
of cell extract is used for each assay. For each assay the
following are mixed together in an 1.5 ml Eppendorf tube and
35 incubated at 37°C for one hour: 25 $\mu$l of cell extract, 5 $\mu$l of
4 mM acetyl coenzyme A, 18 $\mu$l $H_2O$ and 1 $\mu$l ^{14}C-chloramphenicol,
40-60 mCi/mM. After incubation, chloramphenicol (acetylated

and nonacetylated forms) is extracted with ethyl acetate and evaporated to dryness. Samples are resuspended in 25 μl of ethyl acetate, spotted onto a TLC plate and chromatographed in chloroform:methanol (19:1). Chromatographs are analyzed by
5    autoradiography. Spots corresponding to acetylated and nonacetylated $^{14}$C-chloramphenicol are excised from the TLC plate and counted by liquid scintillation for quantitation of CAT activity. Peptide nucleic acids that depress CAT activity in a dose dependent fashion are considered positives.

10          B.    Inhibition of HPV E2 Expression
        The assay for inhibition of human papillomavirus (HPV) E2 by peptide nucleic acids is essentially the same as that for BPV-1 E2. For HPV assays appropriate HPVs are co-transfected into either CV-1 or A431 cells with PSV2NEO using the calcium
15   phosphate method described above. Cells which take up DNA are selected for by culturing in media containing the antibiotic G418. G418-resistant cells are then analyzed for HPV DNA and RNA. Cells expressing E2 are used as target cells for antisense studies. For each PNA, cells are pretreated as
20   above, transfected with E2RE1CAT, and analyzed for CAT activity as above. Peptide nucleic acids are considered to have a positive effect if they can depress CAT activity in a dose dependent fashion.
        EXAMPLE 12
25   **Triplexing of PNA to nucleic acids - Probing protocols**
        Probing was effected in 100μ buffer (S1: 100 mM NaCl, 1 mM ZnSO$_4$, 50 mM NaAc, pH 4.5; KMnO$_4$/dimethyl sulphate (DMS): 10mM Na-cacodylate, 1mM EDTA, pH 7.0 or as other wise noted) containing about 200 cps $^{32}$P-labeled DNA fragment, 0.5 μg calf
30   thymus DNA and the desired amount of PNA. Following a preincubation for 60 min at 37°C, the probing reagent was added and the incubation was continued at room temperature. The reactions were terminated by the addition of a stop-buffer. The DNA was precipitated by addition of 200 μL 2% KAc in 96%
35   EtOH and was analyzed by electrophoresis in 10% polyacrylamide sequencing gels. Radioactive DNA bands were visualized by

autoradiography using amplifying screens and Agfa curix RPA X-ray films exposed at -70°C.

Probing conditions were: S1: 0.5 U/µl, 5 min. stopped with 3 µl 1 M EDTA; KMnO$_4$: 1mM, 15 sec. stopped with 50 µl 1M
5  6-mercaptoethanol, 1.5 M NaAc, pH 7.0; DMS: 1% DMS, 15 sec., stopped as for KMnO$_4$ probing. Samples probed with DMS or KMnO$_4$ were treated with piperidine (0.5 M, 90°C, 20 min.) prior to gel analysis.

EXAMPLE 13

10  **Enzymatic and chemical probing of the binding of PNA T$_2$CT$_2$CT$_4$-LysNH$_2$ to pT8C2A8G2**

For this example the 264 bp EcorI/PvuII fragment of pT8C2A8G2 that was 3'-$^{32}$P-end-labeled at the EcorI site was used. Probing was effected as per the protocols of Example 16.
15  The results are shown in Figure 10. Lane 1 is a control in S1 buffer without S$_1$ nuclease and without PNA. Lanes 2-5: S1 probing, lanes 6-9: KMnO$_4$ probing and lanes 10-13: DMS probing. The following concentrations of PNA were used: 0 µM (lanes 2, 6 & 10), 0.25 µM (lanes 3, 7 & 11), 2.5 µM (lanes 4, 8 & 12) or
20  25 µM (lanes 5, 9 & 13). Lanes S are A+G sequence markers.

The plasmid pT8C2A8G2 contains two targets for PNA T$_2$CT$_2$CT$_4$ (SEQ ID NO:1) in opposite orientation. Therefore both KMnO$_4$ (thymines of the pyrimidine strand) and DMS (guanines of the purine strand) probing was done on the same DNA fragment.
25  Probing with single strand specific nuclease, S$_1$, verified that upon binding of PNA, the pyrimidine strand but not the purine strand of the PNA targets, became susceptible to attack by S$_1$ (Figure 10, lanes 2-5). Interestingly, however, in the absence of PNA S$_1$ attacks the purine strand (lane 2). This is
30  undoubtedly due to the formation of an intramolecular triple helix (H-DNA), a well characterized feature of such polypurine/pyrimidine mirror repeats. It is also observed that in the presence of PNA, the S$_1$ sensitivity extends into the linker between the two PNA targets. This can be explained if
35  both PNA targets are occupied, thereby forming a combined loop where the DNA region between the targets is single-stranded.

- 39 -

This type of binding further illustrates that PNA can be used for directed double strand DNA cleavage by $S_1$.

The $KMnO_4$ probing of the PNA-dsDNA complex reveals that all thymines of the target are oxidized by $KMnO_4$ (Figure 13,
5   lanes 8-9). Concomitantly (in terms of PNA concentration) with the occurrence of $KMnO_4$ susceptibility, virtually full protection against reaction with DMS is observed at the two guanines of the opposite PNA target (see Figure 10, lanes 12-13). These results show that displacement of the pyrimidine
10  strand of the DNA target is accompanied by protection of the major groove upon binding of PNA supporting the model in which two PNA strands are participating in the PNA-DNA complex forming PNA DNA-DNA triplex.

EXAMPLE 14

15  Chemical probing of the binding of PNA $T_2CT_2CT_4$-LysNH$_2$ to pA8G2

For this example, the 248 bp EcoRI/PvuII fragment of pA8G2 was used that was $^{32}P$-end-labeled (5'-labeling, lanes 1-5) or (3'-labeling, lanes 6-9)) at the EcoRI site. Probing was effected as per the protocols of Example 12. The results are
20  shown in Figure 11 wherein lane 1 is a control in S1 buffer without $S_1$ nuclease and without PNA. Lanes 2-5: $KMnO_4$ probing and lanes 6-9: DMS probing. The following concentrations of PNA were used: $0\mu M$ (lanes 2 & 6), $0.25$ $\mu M$ (lanes 3 & 7), $2.5$ $\mu M$ (lanes 4 & 8) or $25$ $\mu M$ (lanes 5 & 9). Lanes S are A+G sequence
25  markers.

In this example, in order to assure that the results were not an artifact of the dimeric target, we performed a similar experiment to that of Example 13 except a plasmid (pA8G2) having only a single PNA target was used. As is seen in Figure
30  14, virtually identical results to those with the dimeric target of Example 13 were obtained with $KMnO_4$ and DMS probing of the single target.

EXAMPLE 15

Effect of pH and ionic strength of the binding of PNA $T_2CT_2CT_4$-
35  LysNH$_2$ to pT8C2A8G2

For this example the 264 bp EcorI/PvuII fragment of pT8C2A8G2 was used that was $3'-^{32}P$-end-labeled at the EcorI

site. Probing was effected as per the protocols of Example 12.
The results are shown in Figure 12 wherein lanes 1 & 9 are
controls without PNA. Lanes 1-7 are KMnO$_4$ probing while lanes
9-15 are DMS probing. The samples of lanes 2-7 & 9-15
5   contained 5 $\mu$M PNA. The samples of lanes 3, 5, 7, 11, 13 & 15
contained 100 mM NaCl in the 10 mM Na phosphate, 1 mM EDTA
buffer. The pH of the buffer was 7.5 (lanes 2, 3 & 10, 11);
6.5 (lanes 4, 5 & 12, 13) or 5.5 (lanes 6,7 & 14, 15). Lanes
are A+G sequence markers.

10      This example shows that the Hoogsteen hydrogen bonding
of cytosine to guanine requires protonation of N3 of cytosine
and consequently interactions involving C$^+$-G Hoogsteen base
pairing are very sensitive to the acidity of the medium. This
has been studied extensively by others with DNA triple helices,
15   and can also be seen with (PNA)$_2$/DNA triplexes using
oligonucleotides.

     The results presented in Figure 12 further show that the
strand displacement binding of PNA T$_2$CT$_2$CT$_4$ (SEQ ID NO:1) to the
target of pT8C2A8G2 is also sensitive to pH. Very little
20   binding is observed at pH 7.5 (lanes 2 & 10) compared to pH 6.5
(lanes 4 & 12) and 5.5 (lanes 6 & 14). Full consistency is
also seen between the KMnO$_4$ and DMS probing results showing
clear correlation between strand displacement and Hoogsteen
type binding of the PNA. We have previously shown that medium
25   ionic strength (50-100 mM Na$^+$) inhibits strand displacement
binding of PNA to double-stranded DNA. The results presented
in Figure 15 (lanes 3, 5, 7) are fully consistent with this.
There is no reason from physio-chemical considerations that
binding of PNA to DNA as a conventional PNA DNA-DNA triple
30   helix should be strongly salt dependent. Therefore, such
complexes could be envisaged under salt conditions that
disfavor strand displacement. However, we see no evidence that
this be the case since no DMS footprint is observed under
"high-salt" conditions (Figure 12, lanes 11, 13 & 15).

35      From Examples 13, 14 and 15 as well as other of the
Examples of this specification, we conclude that sequence
specific binding of PNA to double-stranded DNA involves PNA

- 41 -

DNA-DNA triplexes employing conventional Hoogsteen and Watson-Crick base pair hydrogen bonding for recognition.

**EXAMPLE 16**

**Electron microscopy**

5      A pA98 plasmid containing an $A_{98}/T_{98}$ (SEQ ID NO:42) insert in the PstI site of the polylinker of a pUC19 plasmid was used. 0.1 μg of the pA98 plasmid linearized by the Cfr10I I restriction enzyme was incubated with 0.15 μg of PNA H-$T_{10}$-LysNH$_2$ at 37 °C for 3 hours in 10 μL of the TE buffer (10 mM of
10    Tris-HCl, 1 mM of Na$_3$EDTA, pH 8.0). The complex was diluted with buffer containing 10 mM of Tris-HCl, 10 mM of NaCl, pH 7.5 to the final DNA concentration of 0.2 - 0.5 μg/mL and absorbed to the surface of glow discharged in tripropylamine carbon grid for 2 minutes as described in Duochet, J., *Ultrastruct. Res.*,
15    **1971** *35*, 147-167. The grid was stained in 0.1-0.5 % water solution of uranyl acetate for 10-15 sec and dried. The sample was shadowed with Pt/C (95/5) and studied in a Philips 400 electron microscope at accelerating voltage of 40 kV. The length of the DNA molecules were measured and a histogram was
20    plotted as described by A.V. Kurakin in Micron Microscopica Acta, 22,213-221 (1991). The length of unwound region was measured by its thicker strand, for which the number of base pairs per unit length was assumed to be the same as for the DNA duplex. For analysis by use of a histogram, the molecules were
25    oriented in such a manner that the left boundary of the loop was closer to the end of the molecule.

      The electron microscopy of the PNA-DNA complex was carried out on the $A_{98}/T_{98}$ (SEQ ID NO:42) target contained within the supercoiled plasmid, pA98, linearized with the
30    Cfr10I restriction enzyme, and challenged with a PNA H-$T_{10}$-LysNH$_2$ (SEQ ID NO:2). Full occupancy of the target results in a strand displacement loop of 90-100 bases was detected as is shown in Figure 13. This histogram plot is illustrated in Figure 14. As is seen in Figure 13, the DNA molecules carry an
35    open region in the form of "an eye". In all cases one of the two strands in the open region was thicker than the other one and has the same thickness as the normal DNA duplex. The

- 42 -

thicker strand corresponds to the A-strand covered by PNA, while the thinner strand corresponds to the displaced T-strand. The positions of two branch points of the loops were 47.1% and 51.2%, which within the error of EM-measurements (0.5%),
5    coincides with the ends of the A98/T98 insert (47.5% and 51.0%). The average size of the loop was 4.1%, *i.e.*, 114 bp, which also agrees well with the length of the insert( the standard error is 17bp). In control experiments carried out without added PNA, the eye-like structures were not observed.
10   EXAMPLE 17
     Unwinding of closed circular DNA with PNA
        Relaxed circular DNA was prepared by treating ordinary supercoiled plasmid DNA with an extract containing DNA topoisomerase I, as described by V.I. Lyamichev, et. al., J.
15   Biomolec. Struct. Dynamics, 3, 327-338 (1985). DNA-PNA complexes were obtained by incubation of 1-2 $\mu$g of DNA in 5-20$\mu$L TE with 0.4 optical units./mL of PNA for 4 hours at 20-22 °C. This corresponded to about 10 times molar excess of PNA to its potential binding sites. Agarose (1.5%) gel
20   electrophoresis was performed in the TAE buffer containing 1 $\mu$g/mL of chloroquine at 10 °C for 15 hours at 1.5 V/cm. Two-dimensional gel electrophoresis was preformed in the first direction (from top to bottom) in the TAE buffer and in the second direction (from left to right) in the same buffer with
25   addition of 1 $\mu$g/mL of chloroquine.
        In this example, the efficiency of PNA T10 incorporating into the DNA duplex and displacing the DNA T-strand in solution was shown by unwinding of closed circular DNA by PNA. The pA98 plasmid (Example 16) was prepared in the form of relaxed
30   circles (rcDNA). It was expected that complexing with PNA would unwind about 10 turns of the DNA duplex. Because of topological constraints (the two strands in rcDNA were closed and therefore topologically linked), this unwinding would make the rcDNA molecules behave as if they were positively
35   supercoiled by 10 superturns. This would manifest itself in an increase of electrophoretic mobility of the complex as compared with control rcDNA preincubated in the same buffer without PNA.

- 43 -

This increase in electrophoretic mobility is shown in Figure 15 that shows that after incubation with PNA in the low-salt TE buffer for 4 hours at 20 °C, virtually all rcDNA molecules significantly increased their mobility and moved as if they were highly supercoiled. Two-dimensional gel electrophoresis gels, Figure 16, shows that these molecules behave actually as positively supercoiled ones. This technique is used to resolve individual topoisomers. Using a sample with a wide distribution of topoisomers as reference, it is clearly seen Figure 16 that on the average 8-10 turns of the double helix are released upon binding of PNA. These results indicate that in the TE buffer PNA occupied all available binding sites on the DNA duplex efficiently displacing the DNA T-strand in the A98/T98 (SEQ ID NO:42) insert. When NaCl is added to the incubation buffer, the degree of conversion of rcDNA into rapidly moving species dramatically decreased within a narrow range between 20 and 40 mM of NaCl as seen in Figure 17, panel A. This emphasizes a strong sensitivity of the PNA incorporation to the salt concentration. However, after being formed at low salt, the complex showed remarkable stability and tolerated increasing salt concentrations up to at least 500 mM of NaCl (Figure 20, panel B, lane 4). The complex could be dissociated by alkali treatment (Figure 20, panel B, lane 5).

**EXAMPLE 18**

**Inhibition of RNA polymerase $T_3$ transcription elongation by PNA**

The complex, between PNA $T_{10}$-LysNH$_2$ (SEQ ID NO:2) (1 or 10 uM) and pAIOKS (pBluescriptKS$^+$ in which the d(A$_{10}$) target is cloned into the BamH1 site, analogous to pT10) (100 ng) cleaved with restriction enzyme PvuII, XbaI, or BamH1, was formed by incubation in 14 $\mu$L 10 mM Tris-HCl, 1 mM EDTA, (pH 7.4) buffer for 60 min. at 37 °C. Subsequently 4 $\mu$L 5x concentrated polymerase buffer (0.2 M Tris-HCl, pH 8.0, 125 mM NaCl, 40 mM MgCl$_2$, 10 mM spermidine) was added together with 15 U $T_3$ RNA polymerase and ATP (10 mM), CTP (10 mM), GTP (10 mM), UTP (1 mM) and $^{32}$P-UTP (0.1 $\mu$Ci), and the incubation was continued for 10 min. Following ethanol precipitation, the RNA was analyzed by electrophoresis in sequencing gels and the radiolabeled

- 44 -

bands were visualized by autoradiography using intensifying screens. The results are shown in Figure 18: Lane 1: pTIOKS was cut with PvuH and no PNA was present. Lane 2: pTIOKS was cut with BamHl, and no PNA was present. Lane 3: pTIOKS was cut

5    with XbaI, no PNA. Lane 4: as in lane 1 but in the presence of PNA (1 uM). Lane 5: as lane 3 + PNA (luM). Lane 6: as lane 1 + PNA (10 uM). Lane 7: as lane 3 + PNA (10uM).

As is seen in Figure 18, In the absence of PNA the expected run-off transcript is observed (lanes 1,2 & 3). When

10   the preformed complex between PNA $T_{10}$-LysNH$_2$ (SEQ ID NO:2) and the template was used, a transcript which corresponds in length to transcription arrest at the encounter of the PNA, i.e., one nucleotide longer than the run-off at the BamHl site (lane 3), is produced (lanes 4-7). This shows that transcription

15   elongation by RNA polymerase $T_3$ is effectively blocked by PNA bound to the template strand. If PNA was bound to the non-template strand, no transcript corresponding to blockage by the PNA was observed. Similar results were also obtained using RNA polymerase T$_7$ (data not shown). In this experiment the PNA/DNA

20   complexes were formed in low salt buffer prior to addition of the buffer required for enzyme action since binding of PNA to dsDNA is     inhibited by elevated (> 50 mM Na$^+$) salt concentrations. Once formed, the complexes are exceedingly stable to these conditions.

25   **EXAMPLE 19**

**Inhibition of Taq DNA polymerase primer extension by PNA**

A mixture of 100 500 ng Pvull cleaved plasmid pAIOKS, 0.1 ug M13 reverse primer and 1 ug of PNA $T_{10}$-LysNH$_2$ (SEQ ID NO:2) in 10 ul buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM

30   MgCl$_2$, 0.1 mg/ml gelatine) was incubated for 5 min at 90 °C and then for 60 min at 37 °C. Subsequently, 1 U of Taq polymerase, 1ul dCTP (100 uM), dGTP (100 uM), dTTP (100 uM) and $^{32}$P-dATP (1 uCi) were added and the incubation continued for 5 min at 20 °C. Following the addition of 2 ul dATP, dCTP, dGTP, dTTP

35   (1 mM each), the sample was incubated for 15 min at 60 °C. The DNA was precipitated with ethanol and the samples treated as described for Example 18. The results are shown in Figure 19:

Lane 1: pTIOKS was cut with <u>Pvu</u>II and no PNA was present. Lane 2: same as lane one with PNA present. Lane 3: no PNA and pTIOKS was cut with <u>Bam</u>H1.

As is shown in this example, in a primer extension experiment using a pre-formed complex between PNA $T_{10}$-LysNH$_2$ (SEQ ID NO:2), a single-stranded target, and a primer downstream from the target sequence, a product corresponding in length to blockage at the PNA was produced by <u>Taq</u> DNA polymerase (Figure 19, lane 2). No product was detected in the absence of PNA (lane 3). Similar results were obtained using the large fragment (Klenow fragment) of E. <u>Coli</u> DNA polymerase (data not shown). Thus demonstrates elongation by DNA polymerases is blocked by PNA.

**EXAMPLE 20**

**Inhibition of transcription by PNA**

The PNA oligomers $T_{10}$-LysNH$_2$ (SEQ ID NO:2), $T_5CT_4$-LysNH$_2$ (SEQ ID NO:6) and $T_2CT_2CT_4$-LysNH$_2$ (SEQ ID NO:1) were synthesized as described in Example 1. Plasmids containing the target sequences were obtained by cloning of the appropriate oligonucleotides into the vector pBluescripKS$^+$. To obtain pT10KS and pA10KS, 16-mers 5'-TCGACT$_4$CT$_5$G (SEQ ID NO:36) and 5'-GATCCA$_{10}$G (SEQ ID NO:34) were cloned into the <u>Bam</u>H1 site, and clones containing the insert in either orientation were isolated. pT9CA9GKS was obtained by cloning 5'-TCGACT$_4$CT$_5$G (SEQ ID NO:36) and 5'-TCGACA$_5$GA$_4$G (SEQ ID NO:37) into the <u>Sal</u>I site and pT8C2KS and pA8G2KS were obtained by cloning 5'-GT$_4$CT$_2$CT$_2$CTGCA (SEQ ID NO:39) AND 5'-GA$_2$GA$_2$GA$_4$CTGCA (SEQ ID NO:38) into the <u>Pst</u>I site. E. <u>coli</u> JM103 was used as host in all cases, and transformations and isolation of clones were done by standard techniques. Plasmids were purified by buoyant density centrifugation in CsCl gradients and characterized by dideoxy sequencing.

**Inhibition of RNA polymerase transcription elongation by PNA**

The complex between the desired PNA and the desired DNA (100 ng) cleaved with the desired restriction enzyme was formed by incubation in 14 $\mu$L 10 mM Tris-HCl, 1 mM EDTA, (pH 7.4) buffer for 60 min at 37 °C. Subsequently 4 $\mu$L 5x concentrated

polymerase buffer (0.2 M Tris-HCl, pH 8.0, 125 mM NaCl, 40 mM
$MgCl_2$ 10 mM spermidine) was added together with 15 U of RNA
polymerase and ATP (10 mM), CTP (10 mM), UTP (1 mM) and $^{32}P$-UTP
(0.1 $\mu$Ci). The incubation was continued for 10 min. Following
5   ethanol precipitation, the RNA was analyzed by electrophoresis
in polyacrylamide sequencing gels, and radiolabeled RNA bands
were visualized by autoradiography (using Agfa Curix RPI X-ray
films and intensifying screens). Quantitation was performed by
densitometric scanning using a Molecular Dynamics laser scanner
10  and the ImageQuant™ software.

    Figure 20 illustrates certain concentration dependence
between complexes of pA10KS cleaved with XbaI and the following
concentrations of PNA $T_{10}$-LsyNH$_2$ (SEQ ID NO:2): 0, 0.2, 0.35,
0.7, 1, 1.4, 1.75, 2.1, 2.8, 3.5, 4.3 or 5 $\mu$M (lanes 1-12).
15  The complexes were formed by incubation in TE buffer for 60
min. at 37°C. The buffer was adjusted to transcription
conditions and following the addition of NTPs, $^{32}P$-UTP and $T_3$
RNA polymerase, transcription was allowed to proceed for 5 min.
at 37°C. The transcripts were analyzed by electrophoresis in
20  10% polyacrylamide/7M urea gels and visualized by
autoradiography.

    The results shown in Figure 21 are similar to those of
Figure 20 except that the plasmid pA8H2KS was cleaved with
BamH1 and PNA $T_2CT_2CT_4$-LysNH$_2$ (SEQ ID NO:1) at the following
25  concentrations: 0, 0.6, 1.2, 2.4, 3.6, 4.8, 6, 7.2, 9.6, 14.4
or 16.7 $\mu$M (lanes 1-12).

    Figure 22 illustrates a quantitative representation of
the results shown in Figures 20 and 21. In this figure the
following symbols, respectively, are utilized, squares: pA10KS
30  x XbaI and PNA $T_{10}$-LysNH$_2$ (SEQ ID NO:2); circles: pA8GKS x BamH1
and PNA $T_2CT_2CT_4$-LysNH$_2$ (SEQ ID NO:1). closed squares: relative
amount of the truncated transcript; closed circles: total
amount of transcript.

    This examples illustrates that the phage $T_3$ and $T_7$RNA
35  polymerases can be used as a model system for very efficient
and robust transcription. In this example, a template was
constructed by cloning the appropriate PNA targets into the

polylinker of the BluescriptKS⁺ plasmid.  The target was clone
into the <u>Bam</u>H1 site in both orientations, thus allowing for
four transcription experiments to be performed using either the
constructs and either the $T_3$- or the $T_7$-promoter.  The results

5  presented in Figures 20, 21 and 22 show the dose dependent
inhibition of the transcription of pA10KS (Figure 20) or
pA8G2KS (Figure 21) by PNA $T_{10}$-LysNH$_2$ (SEQ ID NO:2) or $T_2CT_2CT_4$-
LysNH$_2$ (SEQ ID NO:1) using $T_3$ RNA polymerase.  It is observed
that as the concentration is raised the amount of full length

10  transcript is decreased and the amount of truncated product is
increased suggesting that sequence specific transcription
arrest is taking place.  However, the total amount of
transcript also decreases indicating that general inhibition of
transcription also occurs.

15

EXAMPLE 21
Sequence specificity of the transcription elongation arrest by
PNA $T_{10}$-LysNH$_2$ and PNA $T_2CT_2CT_4$-LysNH$_2$ using $T_3$ or $T_7$ RNA
polymerase

20      Samples identical to those used for Example 20 were
analyzed by electrophoresis in a polyacrylamide sequencing gel.
The results are shown in Figure 23 where,  Lanes 1-4: pA10KS/$T_3$
RNA polymerase; lanes 5-8: pT10KS/$T_7$ RNA polymerase; lanes 9-
12: pA8G2KS/$T_3$ RNA polymerase; lanes 13-16: pT8C2KST$_7$ RNA

25  polymerase.  The plasmid used in the samples of lanes 1-3 was
cleaved with Xbal, that in lanes 6-8 with Pst1, that in lanes
9-11 with BamH1 and that in lanes 14-16 with HindIII.  Lanes 1,
8, 9 & 16 are controls without PNA.  Lanes 4, 5, 12 & 13 are
controls cleaved proximal to the PNA target site *i.e.*, with

30  BamH1 (lanes 4 & 5) or Pst1 (lanes 12 & 13).  Samples of lanes
2, 3, 6 & 7 were preincubated with PNA $T_{10}$-LysNH$_2$ (SEQ ID NO:2),
and those of lanes 10, 11, 14 & 15 were preincubated with PNA
$T_2CT_2CT_4$-LysNH$_2$ (SEQ ID NO:1).

      In this example, in order to determine the position of

35  transcription arrest, RNA transcripts were analyzed on high
resolution sequencing gels.  The results show that a specific
RNA transcript is produced by $T_3$-RNA polymerase in the presence

of PNA-T$_{10}$ (SEQ ID NO:2) that is one nucleotide longer than the run-off transcript produced using a template (BamH1 cleaved) ending one nucleotide in front of the PNA target (Figure 23, lanes 2-4). Thus transcription proceeds to, but does not

5    include, the first nucleotide involved in hydrogen bonding with the PNA. The results obtained with T$_7$-RNA polymerase are similar to those with T$_3$-RNA polymerase (Figure 23, lanes 5-8). Interestingly, however, in this case the transcript is less homogeneous in length than the transcript obtained with the T$_3$-

10   polymerase, and the lengths of the transcripts indicate that the polymerase is able to transcribe into the PNA target. These results show that this polymerase is able to some extent to displace the bound PNA during transcription, although the polymerase-PNA encounter may sometimes result in chain

15   termination. Using mixed A/G PNA targets gave the results shown in Figure 23, lanes 9-16. It is noteworthy that transcription arrest with an A$_4$GA$_5$ (SEQ ID NO:43) target and a T$_4$CT$_5$-LysNH$_2$ (SEQ ID NO:8) PNA occurs one to two nucleotides inside the target, while transcription arrest with an A$_4$GA$_2$GA$_2$

20   (SEQ ID NO:44) target and a T$_4$CT$_2$CT$_2$-LysNH$_2$ (SEQ ID NO:9) PNA is much less efficient (as compared to the results of Example 22) and occurs at the end of the target (Figure 23, lanes 10, 11, 14 & 15).

**EXAMPLE 22**

25   **Transcription arrest in pT9CA9GKS plasmid**

In Figure 24 the effect of PNA T$_{10}$-LysNH$_2$ (SEQ ID NO:2) (lanes 1, 2, 4 & 5) on T$_7$ RNA polymerase transcription of pT10KS (lanes 1-3) and pA10KS (lanes 4 & 5), and of PNA T$_5$CT$_4$-LysNH$_2$ (SEQ ID NO:6) (lanes 6, 7, 10 & 11) on T$_3$ or T$_7$ RNA

30   polymerase transcription of pT9CA9GKS (lanes 6-11) are shown. The plasmid of the samples in lanes 1, 2, 4-7, 10 & 11 were cut with PvuII while BamH1 was used for the sample of lane 3, and SalI was used for the samples of lanes 8 & 9. The PNA concentration was 5 $\mu$M.

35   As is shown in this example, experiments using the pT9CA9GKS plasmid and T$_3$ or T$_7$ RNA polymerase showed (Figure 26, lanes 5-11) that transcription arrest at this PNA target

- 49 -

occurred 2-3 bases inside the target.   Employing a gel-
retardation assay, it was ascertained that the complexes
between PNAs $T_{10}$-LysNH$_2$ (SEQ ID NO:2), $T_4CT_5$-LysNH$_2$ (SEQ ID NO:8)
or $T_4CT_2CT_2$-LysNH$_2$ (SEQ ID NO:9) with their complimentary dsDNA
5   targets are stable for at least 15 min when transferred from
the TE-buffer to the transcription assay-buffer, thus
eliminating the possibility that the PNA was dissociating
spontaneously during transcription.   These results indicate
that an increase in the G-content of the target, and thus in
10  the C-content of the PNA decreases the efficiency of
transcriptional arrest due to read-through.   In all cases it is
also observed that PNAs at higher concentrations result in an
unspecific inhibition of transcription by $T_3$ or $T_7$ RNA
polymerase.   We cannot say if this effect is due to an
15  inhibitory effect of the PNA directly at the polymerase, or if
it is caused by binding of the PNA to other sites on the DNA
template, *e.g.*, the promoter.

EXAMPLE 23

Sequence specificity of the transcription elongation arrest by
20  PNA $T_{10}$-LysNH$_2$

In this example, template versus non-template binding of
the PNA was examined.   The experiments were performed
essentially as described for Example 21.   The results are shown
in Figure 25 using plasmid pA10KS (a & c) or pT10KS (b) cleaved
25  with PvuII (lanes 1 & 4), Xbal (lanes 2 & 5) or BamH1 (lane 3).
PNA $T_{10}$-LysNH$_2$ (SEQ ID NO:2) was present at 3.3 $\mu$M in the
samples of lanes 4 & 5.   Panel c is a longer exposure of panel
a.

As is shown in Figure 25, if the PNA is bound to the non-
30  template strand transcription is virtually not arrested (Figure
25, Figure 24: lanes 1-5).   A very weak band corresponding to
arrest at the end of the PNA target can be detected upon longer
exposure of the autoradiogram (Figure 25, panel c, lanes 4,5).
These findings are similar to results reported with DNA
35  templates site specifically modified with covalent psoralen
adducts.

EXAMPLE 24

- 50 -

Sequence specificity of the transcription elongation arrest by
PNA T$_6$, T$_7$ or T$_{10}$-LysNH$_2$ using pA10KS/T$_3$ RNA polymerase

In this example transcription arrest with PNAs shorter
than 10 mers was examined. For the results shown in Figure 26,
5    the plasmid was cleaved with XbaI in all cases and preincubated
with PNA T$_6$-LysNH$_2$ (SEQ ID NO:7) ($\mu$M, lane 2), T$_7$-LysNH$_2$ (SEQ ID
NO:45) ($\mu$M, lane 3) or T$_{10}$-LysNH$_2$ (SEQ ID NO:2) ($\mu$M, lane 4).
Lane 1 is a control without PNA.

The results seen in Figure 26 show that an 8-mer and even
10   a 6-mer PNA, although less efficiently, are able to arrest
transcription by T$_3$-RNA polymerase. The results also show that
the arrest is occurring with the PNA bound at the far end of
the target indicating that the PNA is binding to the 10-mer
target in a floating mode and that the RNA polymerase is
15   "pushing" the PNA.


EXAMPLE 25

Transcription initiation from PNA-DNA displacement loops

Peptide nucleic acid (PNA) form (PNA)$_2$/DNA,DNA triplex-D-
20   loop structure upon binding to complimentary double-stranded
DNA as shown schematically in Figure 27. If two adjacent PNA
sites are present in cis or in trans (wherein cis is where two
PNA bind adjacent to each other on the same strand and trans is
where the two PNA bind on opposite strands), structures of the
25   type shown in Figure 27 b,c are formed. In this example, since
these structures resemble transcription elongation loops we
tested if E. coli RNA polymerase is able to recognize and bind
to such PNA/DNA loops.

The three plasmids pT9C, pT9CT9C (pUC19 derivatives) and
30   pT9CA9GKS (Bluescript derivative) were used for this in vitro
transcription experiments. Restriction fragments were isolated
by digestion with PvuII and on polyacrylamide gels resulting in
the fragments shown in Figure 28. In addition, the isolated
PvuII fragment from pKST9C was restricted with XbaI or SacI
35   before PNA hybridization to obtain a shorter transcript from
the 'PNA-promoter'. PNA-DNA complexes were formed by combining
0.3, $\mu$ PNA (50 OD) with DNA fragments in 10 mM Tris-HCl pH 8.0

- 51 -

and 0.1 mM EDTA in a total volume of 25 $\mu$L for 1 hour at 37°C. The transcription were initiated by addition of 50-100 mM <u>E. coli</u> RNA Polymerase holoenzyme, T3 and T7. The reaction mix contained a final concentration of: 40 mM Tris-HCl pH 8.0, 120

5    mM KCl, 5mM MgCl$_2$, 0.1 mM DTT and 1mM of ATP, CTP, GTP and 0,1, mM of UTP and $^{32}$P UTP. The complete mixture of 30 ul were incubated at 37° for 20 minutes followed by ethanol precipitation. The RNA transcripts were analyzed on 8% denaturing polyacrylamide gels, and visualized by

10   autoradiography. The RNAse H experiment was done by hybridization of a complimentary oligonucleotide to the mRNA to synthesize a double-stranded target for RNAse H.

EXAMPLE 26

Cis competition between the PNA promoter and lacUV5

15      A PvuII fragment including a single or a triple PNA binding site together with the <u>lac</u>UV5 promoter was isolated. The DNA was incubated with increasing amounts of PNA for 1 hour. The transcriptions were performed as described in Example 25. As indicated by the gel-shifts shown in Figure 29,

20   a slower migrating complex between the target DNA and the RNA polymerase is formed only if a PNA binding to this target is formed is prebound to the DNA.

EXAMPLE 27

RNA polymerase footprinting

25      In this example DNase I footprinting experiments are undertaken to define where on the DNA fragment binding takes place. DNA fragments of the length shown in Figure 30 are labeled in the 3' end with klenow polymerase and $^{32}$P-dXTP. The PNA-DNA complexes are formed as described in Example 26 above.

30   The RNA Polymerase-PNA/DNA complexes are formed in a reaction buffer as described for Example 26 with the addition of 2ug/ml calf Thymus DNA in a total volume of 100 ul. After 15 min. incubation at 37 °C the samples are digested with 0.03 $\mu$L DNAse (1mg/ml) for 3 min followed by ethanol precipitation and

35   analyzing by 8% denaturing PAGE. Expected results of this example are also shown in Figure 29. In the presence of PNA but in the absence of RNA polymerase only a weak footprint

- 52 -

corresponding to PNA binding to the target is observed. However, upon addition of E. coli RNA polymerase a clear footprint is observed.

**EXAMPLE 28**

5   **Gel-Shift assay**

PNA-DNA complexes are formed as for the footprint experiment in Example 27. After 15 min at 37 °C incubation the samples are analyzed on 5% PAGE. The gel-shift and DNaseI footprinting results should demonstrate that E. coli RNA

10  polymerase binds to a PNA/dsDNA strand displacement loop. The binding should differ distinctly from that seen for RNA polymerase in the initiation and in the elongation complex. Binding in case of the PNA/DNA complex likely will be confined to the DNA loop and will not extend much into the surrounding

15  double-stranded DNA.

Although the (PNA)$_2$/DNA,DNA triplex-D-loop structurally does resemble an RNA transcription elongation loop, it differs in one important aspect; the PNA does not contain a 3'-hydroxyl group to be used as an elongation substrate for RNA polymerase.

20  Nevertheless, PNA dependent transcription is observed from a DNA molecule containing a PNA target. Furthermore, the length of the resulting transcript should corresponds to a run-off transcript initiated at the bound PNA, as exemplified in Figure 27a. The transcription should be more efficient if a double

25  PNA target is used giving rise to a loop of approximately 30 bases (Figure 27b) in the cis configuration, and approximately 16 bases in the trans configuration (Figure 27c). In the latter case, transcripts of two distinct sizes should be produced (Figure 27c), which in length corresponds to

30  initiation at both PNA targets and proceeding in opposite directions (Figure 27c).

Two experiments will be undertaken to estimate the strength of the PNA dependent transcription initiation. In one experiment, both the PNA target and the strong CAP independent

35  UV5 promoter are present on the same DNA fragment. Upon titration with PNA the full run-off transcript from the UV5 promoter is inhibited, while a new transcript corresponding to

- 53 -

transcription arrest at the PNA site appears.  A very faint band of a length corresponding to transcription initiated at the PNA site is also observed.  However, when a DNA fragment containing a triple PNA binding site is used, transcription

5   from this "PNA-promoter" should be able to compete fully with the UV5 promoter.  As the PNA concentration is increased, transcription from the UV5 promoter decreases with concomitant increase in the amounts of two transcripts that correspond in length to the products expected for transcription in either

10  from the "PNA-promoter".  An analogous experiment should be performed in trans using a mixture of a UV5 containing DNA fragment and either of the DNA fragments containing the single or the cis or trans double PNA targets.  The results of these experiments should confirm that a single PNA decamer target is

15  not able to compete with the UV5 promoter whereas both of the dimeric targets competes very efficiently.  While we do not want to be bound by theory, the results should suggest that a single-stranded DNA loop is a major structural determinant for RNA polymerase upon transcription initiation and elongation.

20  **EXAMPLE 29**

**Site specific termination of *in vitro* transcription in Eukaryotic nuclear extracts**

PNA TTTTCTTTT-NH$_2$ (SEQ ID NO:16), TTCCCTTCC-LysNH$_2$ (SEQ ID NO:12), CCCCCCCTTTTTTTTT-NH$_2$ (SEQ ID NO:18), TTCTCTCTCT-NH$_2$

25  (SEQ ID NO:20) and CCTCCTTCCC-NH$_2$ (SEQ ID NO:19) were synthesized as described in Example 1.  Phosphodiester oligonucleotides were synthesized using standard protocols as for example Vickers, T. et. al., *Nucleic Acids Research*, **1991**, *19*, 3359-3368.  Plasmids were constructed with the CMV IE 1

30  promoter driving the transcription of viral DNA sequences.  The plasmid pIE72 is a CMV IE cDNA clone which was constructed as described by Stenberg, M. et. al., *J. Virol.*, **1993**, *64*, 1556-1565.  pCH495 was constructed by ligating the 495 bp SalI/EcoRI fragment from the plasmid pBH10 (see Hahn, B., et. al., *Nature*,

35  **1984**, *312*, 166-169) into the vector pUC-CMV, prepared by digestion with the same two enzymes.  pCIN was constructed by ligating a 463 bp PvuII fragment from the plasmid pSVCC3 (see

- 54 -

Depeto, A.S. and Stenberg, R.M., *J. Virol.*, **1989**, *63*, 1232-
1238), a genomic CMV clone, into the vector pCEP-4 (Invitrogen)
linearized at the PvuII site.   pCINr was constructed in the
same manner, however the insert is cloned in the opposite
5 orientation.   The plasmids are illustrated in Figure 31.   All
plasmids were constructed with the CMV IE1 promoter (open box)
driving the transcription of sequences containing homopurine
sites for targeting with homopyrimidine PNA.   pCH495 contains
a 495 bp SalI/EcoRI fragment from pBH10 (3738 to 4233 of the
10 HIV-1 genome).    The plasmid pCIN contains a 463 bp PvuII
fragment subcloned from the plasmid pSVCC3, a genomic CMV
clone.    pCINr contains the same fragment in the opposite
orientation.   pIE72 is a CMV IE1 cDNA clone.   The filled boxes
represent the transcribed region of each linearized plasmid.
15 The expected transcript length is shown at the 3' end of each.
The sequence of the PNA target site is also shown along with
the position relative to the end of the transcript.
*In vitro transcription.*   100 ng/ul of linearized plasmid was
pre-incubated with PNA at various concentrations in a volume of
20 10 ul for 3 hours at room temperature in KCl buffered with 10
mM HEPES, pH 6.8.   For studies on pH dependence, KCl was
buffered with 10 mM TrisOAc, pH 5.1 or 10 mM Tris-HCl, pH 7.9.
KCl ranged in concentration from 0.5 to 100 mM.   Following the
pre-incubation, 2 ul of the PNA bound plasmid was transcribed
25 in HeLa nuclear extracts (Promega) following the manufacturer's
protocol.    The transcribed RNA's were precipitated then
separated by electrophoresis through 5% denaturing acrylamide
gels, run 2 hours at 15W.   Gels were then dried and exposed to
film.   Bands on the exposed film were quantitated using a
30 Molecular Dynamics laser densitometer.
        A PNA with the sequence TTTTCTTTT-NH$_2$ (SEQ ID NO:16) and
a non homologous PNA control were pre-incubated with the 200 ng
of EcoRI linerized plasmid pCH495 at the indicated
concentrations for 3 hours at 22 °C in 10 ul of 10 mM KCl, pH
35 6.8 (potassium salt was used in all binding buffers because it
is the most compatable with the transcription extract).
Following the pre-incubation, the PNA bound plasmid was

transcribed in HeLa cell nuclear lystates (Promega) containing
$^{32}P$ GTP. The transcribed RNA product was then electrophoresed
on a 5% acrylamide gel with 50% w/v urea, which was then dried
and exposed to film. The expected size of the full length run

5  off transcript and the PNA truncated transcript are shown in
Figure 31. The concentration at which 50% of the transcript
was truncated was determined to be approximately 80 uM by
quantitating the bands from the exposed film using a Molecular
Dynamics densitometer.      Results   are   shown   in  Figure  32.

10 Linearization of the plasmid with EcoRI and PvuII provides
transcription templates with expected run-off transcripts of
495 (lane1) and 190 (lane 2) nucleotides respectively. In the
presence of the complementary PNA, the EcoRI linearized plasmid
yields two distinct transcripts, one corresponding to the run-

15 off  transcript  and  a  second  truncated  transcript  with  a
molecular weight consistent with the 373 nucleotide product
expected if truncation occurred at the PNA binding site. The
relative proportion of the truncated nucleotide product
increases with increasing concentration of PNA in the pre-

20 incubation.    When a DNA oligonucleotide of the same sequence
(5'-TTTTCTTTT-3') was used in place of the PNA no truncated
product was observed (data not shown). The experiment was
repeated for the other PNA target sites shown in Figure 31.
These results are shown in Table I and are given as $IVT_{50}$; the

25 PNA  concentration  at  which  50%  of  the  transcript  is
specifically truncated. All pre-incubations were carried in 10
mM KCl at pH 5.1, 6.8, or 7.9. In all cases, the amount of
transcript, both full length and truncated, decreases with
increasing PNA concentration, suggesting a non-specific

30 inhibition of transcription by PNA. This effect seems to be
become more pronounced as the percentage of cytosine residues
in the PNA increases. For example, at a PNA concentration of
1 mM, the total yield of transcript is reduced approximately 63
percent in the presence of PNA $TTTTCTTTT-NH_2$ (SEQ ID NO:16), 84

35 percent in the presence of $TTCTCTCTCT-NH_2$ (SEQ ID NO:20) and
completely in the presence of $CCCCCCCTTTTTTTTTT-NH_2$ (SEQ ID
NO:18) (compare the 1 mM lanes in Figure 32 and Figure 33.

- 56 -

Direct addition of PNA to transcription extracts did not result in the production of any truncated product (data not shown).

## TABLE I

**Summary of *in vitro* transcription activity for PNA oligomers.**

| PNA | plasmid | pH 5.1 | pH 6.8 | pH 7.9 |
|---|---|---|---|---|
| TTTTCTTTT-NH$_2$ | pCH495 | 58 uM | 76 uM | >500 uM |
| TTCTCTCTCT-NH$_2$ | pIE72 | 29 uM | 103 uM | >1 mM |
| CCCCCCCTTTTTTTTT-NH$_2$ | pCIN,pCINr | 30 uM | 124 uM | >1 mM |
| CCTCCTTCCC-NH$_2$ | pIE72 | >1 mM | none | none |

EXAMPLE 30

Effect of target site orientation on PNA mediated termination of *in vitro* transcription.

Plasmids pCIN and pCINr were linearized with HindIII, then pre-incubated with PNA of the sequence CCCCCCCTTTTTTTTT-NH$_2$ (SEQ ID NO:18) at the concentrations shown in Figure 33 under the conditions described for Example 29. Following the IVT a truncated product was clearly observed for pCIN, where the PNA is bound to the template strand. However, for pCINr, where the PNA is bound to the non-transcribed strand, only a slight amount of truncated product of the expected size is observed. The plasmids pCIN and PCINr contain a CMV gene fragment cloned in opposite orientations under the transcriptional control of of the CMV promoter. pCIN contains the binding site for the PNA CCCCCCCTTTTTTTTT-NH$_2$ (SEQ ID NO:18) on the template strand, while pCINr contains the same target site on the non-transcribed strand. The PNA was tested for the ability to effect *in vitro* transcription as per Example 29. The results are shown in Figure 33. Both plasmids show non-specific inhibition of *in vitro* transcription at the highest concentration of PNA tested (1 mM). However, when the PNA concentration is decreased to 100 uM, approximately one half of the RNA produced corresponds to the size expected for the PNA truncated product (209 nucleotides) with the pCIN

template.   In contrast, pCINr shows only a faint truncated
product of the correct size at the same concentration.

**EXAMPLE 31**

**Effect of ionic strength and pH on PNA mediated termination of**
5   **transcription.**

25 uM PNA TTTTCTTTT-NH$_2$ (SEQ ID NO:16) was pre-incubated
three hours with pCH495 at pH 6.8 in 0.5 to 100 mM KCl buffer.
Following the pre-incubation the linearized plasmid was
transcribed as detailed in Example 29.   The results are shown
10   in Figure 34 with the percent PNA truncated product indicated
at the bottom of each lane.   The proportion of the truncated
product relative to full length transcript is increased as the
salt concentration is decreased.

The same PNA and plasmid template were used to determine
15   the effects of pH on the ability of PNA to effect
transcription.   PNA at a concentration of 10 or 100 uM, was
pre-incubated (as described in Example 29) with the template in
10 mM KCl buffer at pH 5.1, 6.8, or 7.9 then transcribed in
HeLa cell extracts.   The resultant transcription products are
20   shown in Figure 35 with the percent truncated product shown
below each lane.   At the highest pH tested, 7.9, little
specifically truncated product is observed even at a PNA
concentration of 100 uM.   A slight decrease in pH to 6.8
results in a dramatic increase in the amount of truncated
25   message produced at the 100 uM PNA concentration, although no
truncated message is observed at a PNA concentration of 10 uM.
However, when the pH is decreased to 5.1, a small amount of the
truncated product is observed even at the lower PNA
concentration and is increased at the higher PNA concentration.

30   The experiments were also performed with the other PNA's.
The IVT$_{50}$ for each is listed in Table I, above.   The effect of
pH on PNA binding seems to become more enhanced as the
percentage of cytosine residues in the PNA oligomer increases,
probably due to enhanced binding upon cytosine protonation.   At
35   pH 5.1, one PNA, CCTCCTTCCC-NH$_2$ (SEQ ID NO:19), showed only a
small amount of truncated product at the highest concentration
tested.   At higher pH no truncated product was observed at all.

- 58 -

non-specific effect was generally more pronounced in the oligomers with larger cytosine/thymidine ratios.

The ability of PNA to inhibit transcription *in vitro* supports it use as an antigene therapeutic agent. In using PNA
5   as an antigene agent both pH and ionic barriers would be considered. The effects of pH are mitigated by targeting adenosine rich regions, avoiding dG/dC duplexes. Alternatively, modified residues, such as methyl cytosine could facilitate binding. In this example inhibition of *in vitro*
10  transcription was most efficient at ionic strengths less than those typically found in cells. Therefore, it is possible that PNA will strand invade under physiological conditions if given sufficient time to overcome kinetic barriers. Further to enhance the rate of strand invasion metabolically active
15  regions of the genome would be targeted.

EXAMPLE 32

*In vitro* binding of PNA to target.

The target for *in vitro* binding of duplex DNA was prepared by annealing two 50 base DNA oligonucleotides in which
20  the binding site for PNA TTTTCTTTT-NH$_2$ (SEQ ID NO:16) is centered (5'-AAACAGGGCA GGAAACAGCA TATTTTCTTT TAAAATTAGC AGGAAGATGG-3' and 5'-CCATCTTCC TGCTAATTTT AAAAGAAAAT ATGCTGTTTC CTGCCCTGTTT-3' (SEQ ID NO:46)). The target was $^{32}$P end labeled, then incubated at roughly 500 nM with various concentrations of
25  PNA in the buffers described above for 4 hours at room temperature. PNA bound duplex was separated from free by electrophoresis on a 5% native acrylamide gel in TBE. The gel was then dried and exposed on a Molecular Dynamics PhosphoImager. The K$_d$ is given as the PNA concentration at
30  which one half of the labeled target is bound. Gel shifts of PNA against DNA oligonucleotide complement were carried out in the same buffers using an end labeled DNA oligonucleotide with the sequence 5-AAAAGAAAA-3', which was present in each hybridization mix at 1 nM. Bound DNA oligonucleotide was
35  separated from free by electrophoresis through a 12% native acrylamide gel in TBE.

- 59 -

        Following the above protocols, the *in vitro* binding of
PNA to DNA was assessed by gel mobility shift assay.  The 50
base pair duplex DNA target for the PNA TTTTCTTTT-NH$_2$ (SEQ ID
NO:16) was prepared with the binding site for PNA TTTTCTTTT-NH$_2$
5    (SEQ ID NO:16) centered in the fragment.  The DNA complement
(5'-AAAAGAAAA-3') was synthesized and end labeled.  PNA was
incubated with the either the single strand complement or
duplex DNA target in 20 mM KCl buffer at pH 5.1, 6.8, or 7.9.
PNA bound DNA was then separated from free on a native
10   acrylamide gel.  Figure 36a shows that PNA binding to both the
single stranded and duplex target is effected by the pH of the
binding buffer.  Binding affinity (K$_d$, the PNA concentration at
which one half of the target is shifted in mobility) to both
single stranded and duplex DNA complement was determined by the
15   gel shift assay.  As the pH is decreased affinity increases.
PNA affinity for the single stranded target is much greater
than for the duplex target, presumably due to the stability of
the DNA duplex in 20 mM KCl.  Other experiments to determine
the binding affinity of other PNA oligomers showed that the
20   effect of pH on binding was more pronounced with cytosine rich
PNA's (data not shown).

        Binding of PNA to target was also utilized to study the
effects of salt concentration on PNA binding by incubating PNA
TTTTCTTTT-NH$_2$ (SEQ ID NO:16) with it's single stranded or
25   duplex target in buffers containing 1, 10, or 100 mM KCl at pH
6.8.  The results are shown in Figure 36b.  While the salt
concentration does have a large effect on the ability of PNA to
bind duplex DNA, there is little salt effect on the binding to
single stranded complement.  Similar results were obtained when
30   PNA TTCCCTTCC-LysNH$_2$ (SEQ ID NO:12) was tested (data not
     shown).

**EXAMPLE 33**

**PNA-DNA base pair recognition**

        To test if PNA containing all four natural nucleobases
35   is a true DNA mimic in terms of base pair specific
hybridization to complementary oligonucleotides a pentadecamer
PNA was designed to contain an almost equal number of

- 60 -

hybridization to complementary oligonucleotides a pentadecamer PNA was designed to contain an almost equal number of pyrimidines and purines yet having no more than two purines or pyrimidines juxtaposed. It contains an equal number of
5   thymines and cytosines and a single guanine. Furthermore, the sequence of the pentadecamer is non-selfcomplementary and contains a GTCA sequence at the center. By measuring the thermal stability of complexes between the PNA pentadecamer and the Watson-Crick complementary oligonucleotide, as well as 12
10  other oligonucleotides each having a single base mismatch at one of the four center PNA nucleobases, information about PNA-DNA base pair recognition was obtained. Finally, the PNA was designed with no pyrimidine stretches so as to strongly disfavor - if not prohibit - triplex formation.
15      The thermal stability measured as the melting temperature, $T_m$, of complexes between the pentadecamer, H-TGTACGTCACAACTA-NH$_2$ (SEQ ID NO:10) and the complementary deoxyoligonucleotide 3'-ACATGCAGTGTTGAT (SEQ ID NO:47) (termed anti-parallel orientation: amino terminal of the PNA
20  complementary to the 3'-end of the oligonucleotide) was 69.5°C (see Table II), whereas the $T_m$ of the corresponding PNA-DNA complex in parallel orientation was 56.1°C. The $T_m$ for the PNA-RNA complexes was 72.3 and 51.2°C, respectively (see Table II). The orientation preference was further settled by using
25  two decamer PNAs, and hybridizing these to complementary oligonucleotides in both orientations. The anti-parallel orientation was preferred in all cases (see Table II). Furthermore, it is noteworthy that virtually identical $T_m$s were obtained regardless of which strand is the PNA and which is the
30  DNA (see Table II, 2nd & 3rd row). The results also show that the presence of a terminal lysine amide (add to certain PNAs to reduce aggregation of oligo thymine PNA) does not influence the preferred orientation of the PNA relative to the DNA.

- 61 -

## TABLE II

### Melting temperatures Tₘ(°C) for PNA/DNA, PNA/RNA, DNA/DNA and DNA/RNA complexes.ᵃ

| 1st strand sequenceᵇ | TGTACGTCACAACTAᶜ | GTAGATCACTᵈ | AGTCATCTACᵈ |
|---|---|---|---|
| DNA:DNA | 53.3 | 33.5 | 33.5 |
| DNA:RNA | 50.6 | nd | nd |
| PNA:DNA (parallel) | 56.1 | 38.0 | 38.0 |
| PNA:DNA (anti-parallel) | 69.5 | 51.0 | 49.0 |
| PNA:RNA (parallel) | 51.2 | nd | nd |
| PNA:RNA (anti-parallel) | 72.3 | nd | nd |

a   Absorbance vs. temperature curves were measured at 260
    nm in 100mM NaCl, 10 mM Na-phosphate, 0.1 mM EDTA, pH 7.
    $T_M$, the temperature at which half of the molecules are
    hybridized was obtained by fitting triplicate melting
    curved at 4μM of each strand to a modified two state
    model with linear sloping baseline.
b   Written 5'-3' for oligonucleotides and N- to C- terminal
    for PNA.
c   The PNA terminates in a carboxamide.
d   The PNA terminates in a lysine amide.

When a Watson-Crick base pair mismatch was introduced in
the oligonucleotide at any position facing the four middle PNA
nucleobases (GTCA) in the pentadecamer a large increase in $T_m$
(8-20°C, Figure 37, was observed, thereby providing compelling
evidence that PNA-DNA recognition takes place by Watson-Crick
base pairing, i.e., A-T and G-C base pairing. Qualitatively
similar results were obtained using oligonucleotides in the
parallel orientation. Shown in Figure 37 are the effect of
base pair mismatches on the thermal stability of PNA/DNA
complexes. Thermal stability of complexes between PNA, H-
TGTACGTCACAACTA-NH₂ (SEQ ID NO:10) and the thirteen
oligonucleotides 3'-d(ACATGXYZVGTTGAT), in which
X,Y,Z,V=C,A,G,T for the case where the PNA and DNA sequences
are complementary are shown. In each of the twelve other
oligonucleotides, three of the bases (X,Y,Z,V) were

- 62 -

complementary while the fourth was one of the three non-
complementary nucleobases. For example, when (X=T,Y=G,Z=A)
then V=A or C or G etc.   Thus each of the twelve
oligonucleotides contains one of the twelve possible base pair
5    mismatches relative to the PNA pentadecamer. The $T_m$ of these
complexes are displayed as solid bars in Figure 37.  For
comparison, the results of similar experiments performed with
DNA/DNA duplexes are displayed (hatched bars). Hybridizations
were performed in 10 mM Na-phosphate, pH 7.0, 150 mM NaCl, 1 mM
10   $MgCl_2$.

For comparison we also measured the thermal stabilities
of the corresponding DNA/DNA duplexes as shown in Figure 38.
$\Delta T_m$ values for mismatches are shown in Figure 38.
Hybridizations were as for Figure 37.

15       It is noteworthy that for virtually all base pair
mismatched, the decrease in stability is greater for the
PNA/DNA complex than for the DNA/DNA complex (see Figure 38),
thereby indicating that the sequence discrimination is, if
anything, more efficient for PNA recognizing DNA than for DNA
20   recognizing DNA.

The unambiguous evidence for Watson-Crick base pairing
suggest that these PNA/DNA complexes are duplexes rather than
the $(PNA)_2$/DNA triplexes previously observed with homo-
pyrimidine PNA.  This conclusion was confirmed by titration
25   experiments using ^{32}P-labeled oligonucleotides in a gel
retardation assay.  Complete complex formation was observed at
a 1:1 stoichiometry of [PNA] to [DNA] as shown in Figure 39.
Similar results were obtained using circular dichroism for the
detection of complex formation (data not shown).  In this
30   example titration was by gel-shift of the binding of PNA H-
TGTACGTCACAACTA-NH₂ (SEQ ID NO:10) to the 5'-endlabeled
oligonucleotide 3'-d(ACATGCAGTGTTGAT) (SEQ ID NO:47). The
oligonucleotide was labeled with ^{32}P at the 5'-end using
standard techniques[22]. The oligonucleotide (1nmole, 10³ cpm)
35   was incubated with various amounts of PNA (0-5 nmol) in 10 μ 10
mM Tris-HCl, 1mM EDTA, pH 7.4 for 1 hour at 37°C. The samples
were analyzed by electrophoresis in 20% polyacrylamide gels

- 63 -

(TBE buffer = 89mM Tris-borate, pH 8.3, 1mM EDTA) and the radiolabeled DNA visualized by autoradiography. Concentrations of oligonucleotides and PNA were measured photometrically. Similar results were obtained using the complementary
5  oligonucleotide of reversed polarity (5'-d(ACATGCAGTGTTGAT) (SEQ ID NO:52).

**EXAMPLE 34**

**Structure of the PNA/DNA duplex**

        The $T_m$ data (Table II above) show that PNA, in contrast
10  to DNA or RNA, may bind to complementary DNA or RNA in either orientation, although the antiparallel orientation is preferred. Since the PNA backbone is achiral, the orientation does not by itself impose any steric constraints on the winding of the helix, *i.e.*, left or right handed helicity, and computer
15  modelling also indicates that both binding orientations are possible.

        Information on secondary structure may be obtained from circular dichroism (CD) measurements since these are sensitive to the base pair geometry in the helix. Figure 40 shows the CD
20  spectra of the PNA/DNA, PNA/RNA, DNA/DNA and DNA/RNA duplexes. Shown in Figure 40 is the circular dichroism spectra of PNA/DNA (a: antiparallel; b: parallel), PNA/RNA (c: antiparallel; d: parallel), DNA/DNA (e) and DNA/RNA (f) complexes. The complexes were formed by mixing equal molar amounts of the two
25  complements in distilled $H_2O$. Circular dichroism spectra were recorded on a Jasco 700 instrument at room temperature using an optical path of 1 mm. All measurements were averaged ten times and smoothed. All spectra are largely similar suggesting that PNA/DNA and PNA/RNA duplexes are right handed helices with a
30  base pair geometry not drastically different from that found in a B- or an A-DNA helix. However, it is interesting that the CD-spectra, and thus the structure, of PNA/DNA (or PNA/RNA) duplexes are distinctly different when parallel and anti- parallel complexes are compared. The reference DNA/DNA duplex
35  would be expected to adopt a B-like helix while the RNA/DNA duplex would be expected to adopt a more A-like helix, but the CD-spectra of neither of these short duplexes show typical B-

- 64 -

or A-like features. Thus it is not possible to conclude from these results if the PNA/DNA or PNA/RNA helices are preferentially A- or B-like.

5    **EXAMPLE 35**
**Thermodynamics of the PNA/DNA duplex formation**

Thermodynamics parameters for hybridization can be extracted from thermal stability measurements when these are
10    performed at varying concentrations of the complexes, or from the shape of the thermal denaturation curves. Both methods were used to determine $\Delta H^\circ$, $\Delta S^\circ$ and $\Delta G^\circ$ for formation of the PNA/RNA, DNA/RNA, PNA/DNA and DNA/DNA duplexes (Table III). It is remarkable that the decrease in entropy is almost identical
15    for the formation of DNA/DNA and PNA/DNA duplexes and that both reactions are strongly enthalpy driven.

- 65 -

## TABLE III

Thermodynamic parameters for the formation of PNA/DNA,
PNA/RNA, DNA/RNA and DNA/DNA duplexes with the sequence
TGTACGTCACAACTA present in the PNA strand strand[a].

|  | DNA:RNA | PNA:RNA | DNA:DNA | PNA:DNA |
|---|---|---|---|---|
| $\Delta H$ (kcal/mol)[b] | -94.0 | -109.7 | -91.5 | -103.5 |
| $\Delta S$ (eu)[b] | -264.9 | -291.4 | -253.9 | -276.4 |
| $\Delta G_3$[b] | -11.8 | -19.3 | -12.7 | -17.7 |
| $T_M$ ($^\circ$C, 8$\mu$M)[b] | 49.8 | 72.3 | 53.6 | 68.9 |
| $\Delta H$ (kcal/mol)[c] | -128.9 | -128.5 | -105.3 | -106.6 |
| $\Delta S$ (eu)[c] | -372.8 | -345.9 | -296.2 | -285.8 |
| $\Delta G_{37}$ (kcal/mol)[c] | -13.3 | -21.2 | -13.4 | -18.0 |
| $T_M$ ($^\circ$C, 8$\mu$M)[c] | 50.1 | 72.2 | 53.5 | 68.8 |
| $\Delta H$ (kcal/mol)[d] | -111.5 | -119.1 | -98.4 | -105.0 |
| $\Delta S$ (eu)[d] | -318.9 | -318.7 | -275.1 | -281.1 |
| $\Delta G_{37}$ (kcal/mol)[d] | -12.6 | -20.2 | -13.1 | -17.9 |
| $T_M$ ($^\circ$C, 8$\mu$M)[d] | 50.0 | 72.2 | 53.5 | 68.8 |

[a]   Measured in 100mM NaCl, 10 mM Na-phosphate, 0.1. mM
      EDTA, pH 7.0.
[b]   Obtained by fitting melting curves to a modified two
      state model with linear sloping baselines.
[c]   Obtained from linear plots of 1/$T_M$ versus log
      (concentration).
[d]   Temperature independent parameters calculated as the
      average of the two methods described above.

The finding that $\Delta H^\circ$ and $\Delta S^\circ$ are similar for the
formation of PNA/DNA and DNA/DNA duplexes, combined with the
finding that the PNA/DNA and RNA/DNA duplex structures have
similar stacking according to the CD results, indicates that
single stranded PNA must have much the same degree of base
stacking as single stranded DNA, and thus appears to be highly
structured. The kinetics of PNA/RNA duplex formation was also

measured, and the results show that the rate of hybridization is at least as fast as that for 2'O-Me-RNA/RNA or DNA/DNA duplex formation (Table IV). Again fully consistent with the suggestion that the single stranded PNA is at least as

5 prestructured for duplex formation as is DNA (or RNA).


## TABLE IV

### Equilibrium and rate constants for duplexes.

| Duplex | [Na+] (M) | T(°C) | $K_D^a$ | $k_1 (M^{-1}s^{-1})^{a,b}$ |
|---|---|---|---|---|
| PNA/RNA[c] | 0.1 | 37 | $5 \times 10^{-11}$ | $2 \times 10^6$ |
| 2'-O-Me/RNA[c] | 0.1 | 37 | $5 \times 10^{-11}$ | $2 \times 10^5$ |
| DNA/RNA[c] | 0.1 | 37 | $2 \times 10^{-9}$ | - |
| DNA/DNA[d] | 0.2 | 39 | - | $3 \times 10^5$ |

[a]   The buffer was 100 mM NaCl, 10 mM phospphate, 1 mM EDTA, pH 7.0, and the dissociation constant was determined from gel-shift experiments using a $^{32}$P-labeled RNA. The rate was measured in experiments which were stopped by adding 50-fold excesss of unlabeled RNA followed by rapid freezing.

[b]   The rate constants were determined from plots of the pseudo-first order rate constant versus the reagent concentration, which in all cases were much higher than the concentration of the labeled RNA target.

[c]   The sequence of the reagent was: N-(or5')-TGTACGTCACAACTA-C-terminal (or 3'), and the sequence of the RNA target: 5'-TAGTTGTGACGTACA-3'. 2'-O-Me is the 2'-O-methyl RNA derivative.

[d]   Results taken from Tibanyenda et. al., *Eur. J. Biochem.* **1984**, *139*, 19. The sequence was: 5'-CAACTTGATATTAATA.


The following chemical synthesis are shown in Schemes I

35 and II.

### Scheme I

### EXAMPLE 36

### MonoBoc-diethylenetriamine dihydrochloride (1)

A solution of t-butyl-p-nitrophenyl carbonate (10 g;

40 0.0418 mole) in CHCl$_3$ (400 ml) was added to a solution of diethylenetriamine (45 ml; 0.0417 mole) in CHCl$_3$ (250 ml) at 0 °C over a period of 3h. The reaction mixture was stirred overnight at room temperature. The precipitate that appeared was filtered and washed in CHCl$_3$. The solvent was evaporated,

45 first under reduced pressure with a water-aspirator, then with

an oil-pump (0.05 mmHg; 50 °C). The residue was dissolved in a mixture ethylacetate (50 ml)/H$_2$O (50 ml) and the solution was acidified to pH 4 with HCl 4N, extracted with ethylacetate (3x50 ml). The aqueous solution was adjusted to pH 9 with NaOH
5   2N and extracted with ethylacetate (3x50 ml). The aqueous phase was adjusted to pH 11.5 and extracted with ethylacetate (10x50 ml). The combined organic phases of the last extraction were evaporated under reduced pressure and the resulting oil was dissolved in water (50 ml) and acidified to pH 5.
10  Evaporation of water yielded a slightly yellow solid, which was thoroughly washed with ether (yield: 6.41 g; 55%). $^1$H-NMR (D$_2$O): δ (ppm): 1.4 (s, 9H); 3.0 (t, 2H); 3.3 (s broad, 4H); 3.4 (t, 2H)

EXAMPLE 37

15  **Boc-, Z-diethylenetriamine hydrochloride (2)**

        To a solution of 1 (5.5 g; 19.9 mmoles) in dioxane (50 ml)/water (50 ml) adjusted at pH 11 was added a solution of benzylnitrophenyl carbonate (5.44 g; 19.9 moles) in dioxane (50 ml) at 0 °C over a period of 1h, while maintaining the pH at 11
20  with NaOH 2N. The reaction mixture was then stirred at room temperature for 1.5 h. Subsequently, ethylacetate (100 ml) was added and the reaction mixture was cooled at 0 °C and acidified to pH 4 with HCl 4N. A precipitate appeared in the organic phase. The two phases were separated and the organic phase
25  filtered. This (addition of ethylacetate, separating the two phases and filtering the organic phase) was repeated 3 times to give 3.05 g of white solid. An additional 1.38 g was obtained by joining all the organic phases and adding ether (Yield: 4.427 g; 59%). $^1$H-NMR (DMSO d6): δ (ppm): 1.5 (s, 9H); 3.0 (m,
30  4H); 3.4 (m, 4H); 5.1 (s, 2H); 7.4 (s, 5H); 7.1 (m, 1H); 7.6 (m, 1H). MS FAB+ : M+1: 338.2

EXAMPLE 38

**Coupling of 2 with N-1-carboxymethylthymine**

        Compound 2 (4 g; 10.7 mmoles), N-1-carboxymethylthymine
35. (1.97 g; 10.7 mmoles), DhbtOH (1.75 g; 10.7 mmoles) and DIEA (1.2 ml; 12 mmoles) were dissolved in a mixture of CH$_2$Cl$_2$ (50 ml)/DMF (50 ml). After cooling at 0 °C, DCC (2.2 g; 10.7

mmoles) was added and the reaction mixture was stirred at 0 °C for 1/2 h and 2.5 h at room temperature. DCU was filtered and washed with $CH_2Cl_2$ (2x100 ml). The combined organic phases were washed with $NaHCO_3$ 1M (3x100 ml), $KHSO_4$ 1M (3x100 ml), $H_2O$

5  (2x100 ml), dried over $Na_2SO_4$, filtered and evaporated under reduced pressure. The resulting oil was crystallized in chloroform/ether to yield 4.04 g (75 %) of product. $^1$H-NMR ($CDCl_3$); δ (ppm); 1.4 (s, 9H); 1.9 (s, 3H); 3.4 (m, 8H); 4.4 (d, 2H); 5.1 (d, 2H); 5.5 (m, 1H); 5.9 (m, 1H); 6.9 (s, 1H);

10  7.4 (s, 5H); 9.0 (s, 1H). MSFAB+ : M+1; 504.

EXAMPLE 39

Hydrogenolysis of 3 to 4

   Compound 3 (4g; 7.9 mmoles) was dissolved in MeOH (150 ml). At 0 °C, 10% Pd/C (1.4 g) was added. The reaction

15  mixture was hydrogenated at 0 °C for 1h, filtered through Celite and the solvent removed under reduced pressure to give 3.0 g (100%) of 4. $^1$H-NMR ($CDCl_3$); δ (ppm); 1.4 (s, 9H); 1.9 (s, 3H); 3.0-3.8 (m); 4.6 (d, 2H); 6,9 (s, 1H); 7.3 (s, 5H); 9.0 (s, 1H). MSFAB+:M+1: 370.2

20  EXAMPLE 40

N-(benzyl acetate)-glycine ethyl ester hydrochloride (5)

   At room temperature, benzyl bromo acetate (5.7 ml; 36 mmoles) was added over a period of 10 min. to a solution of glycine ethyl ester hydrochloride (5 g; 36 mmoles) and

25  triethylamine (10.43 ml; 0.072 moles) dissolved in absolute EtOH (100 ml). After 4 days at room temperature, the solvent was removed under reduced pressure. The residue was dissolved in ethylacetate(50 ml)/$H_2O$ (25 ml). After separation of the two phases, the organic phase was thoroughly washed with water

30  (8x25 ml). After evaporation of the solvent, the resulting oil was dissolved in ether (20 ml)/water (30 ml) and acidified to pH 4.5. After separation of the two phases, the aqueous phase was concentrated and the resulting oil crystallized in cold ether to yield 4 g (39%) of the product. $^1$H-NMR ($CDCl_3$); δ

35  (ppm); 1.3 (t, 3H); 4.1 (d, 4H); 4.4 (q, 2H); 5.3 (d, 2H); 5.5 (m, 1H); 7.4 (s, 5H). MS FAB+; M+1: 252.1

- 69 -

EXAMPLE 41

Coupling of 5 with N-1-carboxymethylthymine to give 6

   Compound 5 (2.8 g; 9.7 mmoles), N-1-carboxymethylthymine
(1.79 g; 9.7 mmoles), DhbtOH (1.6 g; 9.7 mmoles) and DIEA (1.7
5  ml; 10 mmoles) were dissolved in a mixture of $CH_2Cl_2$ (30 ml)/DMF
(30 ml). After cooling at 0 °C, DCC (2.0 g; 9.7 mmoles) was
added and the reaction mixture was stirred at 0 °C for 1/2 h
and 4h at room temperature. The DCU was filtered and washed
with $CH_2Cl_2$ (120 ml). The combined organic phases were washed
10  with $NaHCO_3$ 1M (3x60 ml), $KHSO_4$ 1M (3x60 ml), $H_2O$ (2x60 ml),
dried over $Na_2SO_4$, filtered and evaporated under reduced
pressure. The resulting oil was crystallized in chloroform (60
ml)/petroleum ether (120 ml) to yield 3.235 g (80%) of 6. [1]H-
NMR (DMSO d6): δ (ppm); 1.3 (dt, 3H); 1.9 (s, 3H); 4.2 (m, 4H);
15  4.4 (d, 2H); 5.3 (d, 2H); 7.2 (s, 1H); 7.4 (s, 5H); 11.3 (s,
1H). MS FAB+: M+1: 418.1

EXAMPLE 42

Hydrogenolysis of 6 to 7

   Compound 6 (4 g; 9.6 mmoles) was dissolved in MeOH
20  (100ml)/DMF (25 ml). At 0 °C, 10% Pd/C (1.6 g) was added. The
reaction mixture was hydrogenated at 0 °C for 1H, filtered
through Celite and the solvent removed under reduced pressure
to give 3.12 g (99%) of 7. ^{1}H-NMR (DMSO d6): δ (ppm); 1.3 (dt,
3H); 1.9 (s, 3H); 4.0-4.5 (m); 4.9 (s, 2H); 7.2 (s, 1H); 11.3
25  (s, 1H). MS FAB+ : M+1: 328.1

EXAMPLE 43

Coupling of 4 and 7 to give 8

   Compound 4 (2.8 g; 7.6 mmoles), 7 (2.5 g; 7.6 mmoles),
DhbtOH (1.24 g; 7.6 mmoles) were dissolved in a mixture of
30  $CH_2Cl_2$ (50 ml)/DMF (50 ml). After cooling at 0 °C, DCC (1.56
g; 7.6 mmoles) was added and the reaction mixture was stirred
at 0 °C for 1/2 h and overnight at room temperature. The DCU
was filtered and washed with $CH_2Cl_2$ (250 ml). The combined
organic phases were washed with $NaHCO_3$ 1M (3x100 ml), $KHSO_4$ 1M
35  (3x100 ml), $H_2O$ (2x100 ml), dried over $Na_2SO_4$, filtered and
evaporated under reduce pressure. The resulting oil was
crystallized in chloroform (30 ml)/petroleum ether (60 ml).

- 70 -

This material was recrystallized in chloroform (30 ml)/ petroleum ether (25 ml) to give 0.608 g (12%) of 8. MS FAB+: M+1: 418.1

**EXAMPLE 44**

5  **Hydrolysis of 8 to give 9**

Compound **8** (25 mg, 0.037 mmoles) was dissolved in absolute ethanol (5ml). Then NaOH 1M (1 ml) was added. The reaction mixture was stirred for 35 min. at room temperature and after cooling, neutralized with Dowex 50W (ca. 1 g). After
10  filtration and evaporation of the solvent, the acid **9** was obtained as a white solid, yield 18 mg (82 %).

**Scheme II**

**EXAMPLE 45**

**Alanine benzyl ester (10)**

15  Thionylchloride (9.8 ml, 135 mmol, 1.2 eqv) was added dropwise during 15 min. to benzyl alcohol (210 ml) stirred under nitrogen and cooled to -10°C. Alanine (10.0 g, 112 mmol, 1.0 eqv) was added in portions during 10 min. The reaction mixture was heated overnight at 60-70 °C. The reaction mixture.
20  was evaporated *in vacuo* and the residue was dissolved in water (150 ml). The pH was adjusted to 1-2 by addition of 4N HCl (aq.) and the aqueous phase was extracted with dichloromethane (2x200 ml). The organic phase was washed with hydrochloric acid (pH 1, 1x50 ml). The aqueous phases were collected,
25  alkalinized (to pH 9-10) by addition of 2N NaOH (aq.) and subsequently extracted with dichloromethane (2x200 ml and 1x100 ml - *alkaline extract*). The alkaline extract was dried ($Na_2SO_4$), filtered and evaporated under reduced pressure affording 13.39 g (67%) of the title compound as a colorless
30  oil. The product was characterized as its hydrochloride: The product (0.50 g, 2.8 mmol) was dissolved in ether (5 ml) and HCl in ether (2ml) was added. The precipitate was collected by filtration and washed with ether (2x5 ml) yielding 0.57 g (95%) of white crystalline alanine benzyl ester hydrochloride. $^1$H-
35  NMR ($D_2O$/TMS): δ 1.48 (unresolved d, 3H, Me); 4.15 (unresolved q, 1H, CH); 5.20 (s, 2H, $CH_2$); 7.38 (br s, 5H, Ph). $^{13}$C-NMR ($D_2O$/TMS): δ 14.7 (Me); 48.4 (CH); 68.1 ($CH_2$); 128.1, 128.5,

134.3 (Ph); 170.1 (C=O). MS(FAB+) m/z (%): 180 (100, M-HCl+H).
Anal. Calcd. for $C_{10}H_{14}ClNO_2$: C, 55.69; H, 6.54; N, 6.49. Found:
C, 55.60; H, 6.56; N, 6.40.

**EXAMPLE 46**

5   N-(2-Boc-aminoethyl) alanine benzyl ester hydrochloride (11)

Alanine benzyl ester (10, 8.82g, 49 mmol, 1 eq) was
dissolved in MeOH (100 ml) and glacial acetic acid (10.34 g,
172 mmol, 3.5 eq) was added. The mixture was stirred at 0 °C
under nitrogen and sodium cyanoborohydride (10.82 g, 172 mmol,
10  3.5 eq) was added. Boc-aminoacetaldehyde (15.67 g, 98 mmol, 2
eq) in MeOH (200 ml) was added dropwise during 2h. The
reaction mixture was stirred at 0 °C for 35 min. and then at 4-
5°C overnight. Water (300 ml) was added to the reaction
mixture and the pH was adjusted to 9 by addition of solid
15  sodium carbonate. The saturated solution was filtered and
subsequently extracted with ether (3x500 ml). The organic
phase was washed with a 1:1 mixture of saturated aqueous
solutions of sodium chloride and sodium bicarbonate (1x450 ml),
dried (Na$_2$SO$_4$), filtered and evaporated under reduced pressure
20  to give 22.60 g of crude N-(2-Boc-amino-ethyl)alanine benzyl
ester as a slightly golden oil. The crude product was
dissolved in dry ether (500 ml) and stirred at 0 °C HCl in
ether (60 ml) was then added to the solution, and the resulting
precipitate was collected by filtration and washed with cold
25  dry ether (3x20 ml) affording 7.21 g (41%) of 11 as white
crystals. A second crop of 2.12 g (12%) could be collected as
a white semicrystalline material by keeping the mother liquor
in the freezer overnight. This second crop was slightly less
pure than the first crop. $^1$H-NMR (D$_2$O/TMS): δ 1.36 (s, 9H,
30  Boc); 1.52 (d, J=7.2 Hz, 3H, Me); 3.16 (m, 2H, NCH$_2$); 3.35 (m,
2H, NCH$_2$); 4.17 (q, J=7.2 Hz, 1H, CH); 5.26 (s, 2H, CH$_2$Ph); 7.40
(s, 5H, Ph). $^{13}$C-NMR (D$_2$O/TMS): δ 13.5 (Me); 27.3 (Box); 36.4
(CH); 45.5 (NCH$_2$); 55.2 (NCH$_2$); 68.4 (CH$_2$Ph); 128.3, 128.7,
128.8 (Ph). MS(FAB+) m/z (%): 323 (100, M-HCl+H) Anal. Calcd
35  for $C_{17}H_{27}ClN_2O_4$: C, 56.90; H, 7.58; N, 7.81; Cl, 9.88. Found :
C, 55.11; H, 7.52; N, 7.93; Cl, 10.45.

- 72 -

EXAMPLE 47

**N-(2-Boc-aminoethyl)-N-(1-thyminylacetyl)alanine benzyl ester (12)**

      *N*-(2-Boc-aminoethyl)alanine benzyl ester hydrochloride
5   (11, 3.00 g, 8.4 mmol, 1.0 eq) was dissolved in dry
dichloromethane (50 ml) and stirred at 0 °C under nitrogen.
DIEA (1.5 ml, 8.4 mmol, 1.0 eq), DHBtOH (1.50 g, 9.2 mmol, 1.1
eq) in dichloromethane (30 ml), 1-thyminylacetic acid (1.69 g,
9.2 mmol, 1.1 eq) in dichloromethane (30 ml) and DCC (2.07 g,
10  10.0 mmol, 1.2 eq) in dichloromethane (40 ml) were added in
that order.  The reaction mixture was stirred at 0 °C for 1 h
and then at 4-5 °C overnight.  The precipitated DCU was
filtered off and washed with dichloromethane (3x15 ml).  The
filtrate was washed with saturated aqueous sodium bicarbonate
15  (3x150 ml) diluted to five times its volume, saturated aqueous
potassium hydrogen sulfate (2x150 ml) diluted to three times
its volume, saturated aqueous sodium chloride (1x150 ml), dried
($Na_2SO_4$), filtered and finally evaporated under reduced pressure
yielding 3.83 g of crude product as a pink glassy foam.  The
20  crude product was chromatographed (silica, 0.063-0.200 mm)
using MeOH/dichloromethane as the eluent (3/97, v/v until the
first fraction appeared then 5/95, v/v) affording 2.87 g (70%)
of **12** a white glassy foam.  **12** was isolated as mixture of two
conformers due to restricted rotation around the amide bond.
25  Consequently, some of the signals in the NMR spectra were split
into a major and minor component.  The values provided below
are for the major component. $^1$H-NMR ($CDCl_3$/TMS): δ 1.44 (s, 9H,
Boc); 1.52 (d, J=7.1 Hz, 3H, ala-Me); 3.31 (m, 2H, $NCH_2$); 3.44
(m, 1H, $NCH_2$); 3.55 (m, 1H, $NCH_2$); 4.34 (q, J=7.3 Hz, 1H, CH);
30  4.53 (s, 2H, acetyl-$CH_2$); 5.16 (s, 2H, C̲H̲$_2$Ph); 6.90 (s, 1H,
thymine-H-6); 7.35 (m, 5H, Ph). $^{13}$C-NMR ($CDCl_3$/TMS): δ 12.2
(thymine-Me); 14.3 (ala-Me); 28.3 (Boc); 39.3 (CH); 46.0
(acetyl-$CH_2$); 48.0 ($NCH_2$); 55.6 ($NCH_2$); 67.2 (C̲H̲2Ph); 79.8
(Boc); 110.5 (thymine-C-5); 128.0, 128.3, 128.5, 128.6, 135.1
35  (Ph); 140.8 (thymine-C-6); 150.9 (thymine-C-2); 155.8 (Boc-
C=O); 164.1 (thymine-C-4); 166.8 (acetyl-$CH_2$); 171.3 (ester-
C=O).  MS(FAB+) m/z (%): 489 (64, M+H). Anal. Calcd for

- 73 -

$C_{24}H_{32}N_4O_7$: C, 59.01; H, 6.60; N, 11.47.    Found: C, 58.55; H, 6.65; N, 11.16.

**EXAMPLE 48**

**N-(2-Boc-aminoethyl)-N-(1-thyminylacetyl)alanine (13)**

5      *N*-(2-Boc-aminoethyl)-N-(1-thyminylacetyl)alanine benzyl ester (**12**, 2.02 g, 4.1 mmol, 1 eq) was dissolved in MeOH (100 ml) and stirred at 0 °C under nitrogen.    10% palladium on activated carbon (1.7 g) was added and the mixture was hydrogenated at atmospheric pressure and 0 °C for 1 h at which 10   time hydrogen uptake had ceased (91 ml, 4.1 mmol, 1 eq had been consumed).    The reaction mixture was filtered through celite which was washed thoroughly with MeOH.    The collected filtrates were evaporated giving 1.65 g (100%) of **13** as a white glassy foam.    **13** was isolated as a mixture of two conformers 15   due   to   restricted   rotation   around   the   amide   bond. Consequently, some of the signals in the NMR spectra were split into a major and minor component.    The values provided below are for the major component. $^1$H-NMR (DMSO-$d_6$/TMS): $\delta$ 1.42 (d, J=7.1 Hz, 3H, ala-Me); 1.46 (s, 9H, Boc); 3.26 (m, 2H, NCH$_2$); 20   3.42 (m, 2H, NCH$_2$); 4.39 (q, J=7.0 Hz, 1H, CH); 4.69 (s, 2H, acetyl-CH$_2$); 7.00 (s, 1H, thymine-H-6); 11.35 (s, 1H, COOH). $^{13}$C-NMR (DMSO-$d_6$/TMS): $\delta$ 12.0 (thymine-Me); 14.8 (ala-Me); 28.3 (Boc); 33.6 (CH); 45.7 (acetyl-CH$_2$); 48.0 (NCH$_2$); 54.8 (NCH$_2$); 78.1 (Boc); 108.1 (thymine-C-5); 142.2 (thymine-C-6); 151.0 25   (thymine-C-2); 155.8 (Boc-C=O); 164.4 (thymine-C-4); 166.9 (acetyl-CH$_2$); 172.6 (acid-C=O). MS(FAB-) m/z (%): 397 (100, M-H); 398 (21, M-H+1); 399 (5, M-H+2). Anal. Calcd for $C_{17}H_{26}N_4O_7$: C, 51.25; H, 6.58; N, 14.06.

**EXAMPLE 49**

30 **Chimera Macromolecule Having Peptide Nucleic Acids Section Attaching to 3′ Terminus of a 2′-Deoxy Phosphorothioate Oligonucleotide Section**

A first section of peptide nucleic acids is prepared as per PCT patent application PCT/EP/01219.    The peptide nucleic 35   acids are prepared from the C terminus towards the N terminus using   monomers   having   protected   amine   groups.    Following completion of the peptide region, the terminal amine blocking

group is removed and the resulting amine reacted with a 3'-C-
(formyl)-2',3'-dideoxy-5'-trityl nucleotide as prepared as per
the procedure of Vasseur, et. al., J. Am. Chem. Soc. 1992, 114,
4006.   The condensation of the amine with the aldehyde moiety

5    of the C-formyl nucleoside is effected as per the conditions of
the Vasseur, ibid., to yield an intermediate oxime linkage.
The  oxime  linkage  is  reduced  under  reductive  alkylation
conditions of Vasseur, ibid., with HCHO/NaBH$_3$CN/AcOH to yield
the nucleoside connected to the peptide nucleic acid via an

10   methyl  alkylated  amine  linkage.    An  internal  2'-deoxy
phosphorothioate nucleotide region is then continued from this
nucleoside as per standard automatated DNA synthetic protocols
(see Oligonucleotide synthesis, a practic approach, M.J. Gait
ed, IRL Press, 1984).

15   EXAMPLE 50

Chimera Macromolecule Having Peptide Nucleic Acids Section
Attaching to 5' Terminus of a Phosphorothioate Oligonucleotide
Section

        A  phosphorothioate  oligonucleotide  is  prepare  in  the

20   standard  manner  on  a  solid  support  as  per  standard  protocols
(see Oligonucleotides and Analogues, A Practical Approach, F.
Eckstein Ed., IRL Press, 1991.   The dimethoxytrityl blocking
group on that nucleotide is removed in the standard manner.
Peptide  synthesis  for  the  peptide  region  is  commenced  by

25   reaction of the carboxyl end of the first peptide nucleic acid
of this region with the 5' hydroxy of the last nucleotide of
the DNA region.   Coupling is effected via DEA in pyridine to
form an ester linkage between the peptide and the nucleoside.
Peptide synthesis is then continued in the manner of patent

30   application PCT/EP/01219 to complete the peptide nucleic acid
region.

EXAMPLE 51

Double-stranded Duplex Structures That Include Chimera Stand
and Triple-stranded Triplex Structures That Include Chimera

35   Strand

        Duplex and triplex structures will be formed with the
chimera strands of Examples 42 and 43 as per the protocols of

other of the above examples.  Duplex structures can include
duplexes between a PNA-RNA or PNA-DNA strand and a RNA strand,
a PNA-RNA or PNA-DNA strand and a DNA strand, a PNA-RNA or PNA-
DNA strand and a PNA strand or a PNA-RNA or PNA-DNA strand and

5   a further chimeric PNA-DNA or PNA-RNA strand.  Triplex
structures can include a PNA containing chimeric strands
triplexing with dsDNA or with a double-stranded PNA construct.
Further triplex structures can include two of the PNA
containing chimera triplexing with a single DNA or RNA strand.

10  Additional triplex structures can include a single PNA
containing chimera plus a PNA triplexing with an additional PNA
containing chimera, an additional PNA strand or a DNA or RNA
strand.

**EXAMPLE 52**

15  **Binding between single strand PNA containing chimera and
transcription factor or other protein**

A PNA containing chimeric strand will be used to bind to
or otherwise modulate single-stranded DNA, double-stranded DNA,
RNA, a transcription factor or other protein.  In use of a PNA

20  containing chimera, part of the binding between the chimera and
the transcription factor or other protein will be binding
between the sugar-phosphate backbone  of the DNA or RNA portion
of the chimera and hydrogen bonding between the ligands, *e.g.*,
nucleobases, of the PNA portion of the chimera.  Binding to the

25  sugar-phosphate backbone includes binding to phosphodiester
linkages, phosphorothioate linkages or other linkgages that may
be used as the bacbone of the DNA or RNA.  In other instances,
bonding can include hydrophobic contacts between hydrophobic
groups on the ligands, including nucleobases, of the PNA or the

30  nucleobases of the nucleic acid portion of the chimera with
like hydrophobic groups on proteins that are being bound.  Such
hydrophobic groups on the chimeric strand include the methyl
groups on thymine nucleobases.

- 76 -

EXAMPLE 53

**PNA Dimer (detT-idaT) Having "Reversed" Monomer**
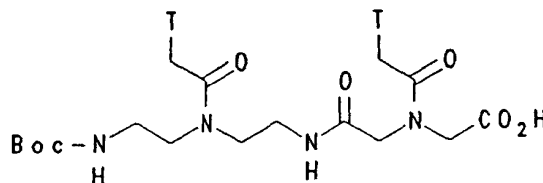


A.    **MonoBoc-diethylenetriamine dihydrochloride.**

A solution of t-butyl-p-nitrophenyl carbonate (10g; 0.0418 mole) in CHCl₃ (400 ml) was added to a solution of diethylenetriamine (45 ml; 0.417 mole) in CHCl₃ (250 ml) at 0°C over a period of 3h. Then, the reaction mixture was stirred overnight at room temperature.

The precipitate that appeared was filtered and washed with CHCl₃. The solvent was evaporated, first under reduced pressure with a water-aspirator, then with an oil-pump (0.05 mmHg; 50°C).

The residue was dissolved in a mixture of ethyl acetate (50ml) and H₂O (50ml). The solution was acidified to pH 4 with 4N HCl and extracted with ethyl acetate (3x50ml). The aqueous solution was adjusted to pH 9 with 2N NaOH and extracted with ethyl acetate (3x50ml). Then, the aqueous phase was adjusted to pH 11.5 and extracted with ethyl acetate (10x50ml). The combined organic phases of the last extraction were evaporated under reduced pressure and the resulting oil was dissolved in water (50ml) and acidified to pH 5. Evaporation of water yielded a slightly yellow solid, which was thoroughly washed with ether (yield: 6.41 g; 55%).

¹H-NMR (D₂O): δ (ppm): 1.4 (s,9H); 3.0 (t,2H); 3.3 (s broad, 4H); 3.4 (t, 2H)

B.    **Boc-, Z-diethylenetriamine hydrochloride.**

To a solution containing the product of Example 53A (5.5 g: 19.9 mmoles) in dioxane (50ml) / water (50ml) adjusted at pH 11 was added a solution of benzyl-nitrophenyl carbonate (5.44g; 19.9 mmoles) in dioxane (50ml) at 0°C over a period of 70 min, while maintaining the pH at 11 with 2N NaOH. The reaction mixture was then stirred at room temperature for 1.5h.

- 77 -

Subsequently, ethyl acetate (100ml) was added and the reaction mixture was cooled at 0°C and acidified to pH 4 with 4N HCl. A precipitate appeared in the organic phase. The two phases were separated and the organic phase filtered.

These operations (adding ethyl acetate, separating the two phases and filtering the organic phase) were repeated 3 times. A white solid (3.05g) was collected this way. An additional 1.38 g was obtained by joining all the organic phases and adding ether. (Yield : 4.427 g; 59%)

$^1$H-NMR (DMSO d6): δ (ppm): 1.5 (s,9H); 3.0 (m,4H); 3.4 (m,4H); 5.1 (s,2H); 7.4 (s,5H); 7.1 (m,1H); 7.6 (m,1H). MS FAB+: M+1: 338.2

C.    Coupling with N-1-carboxymethylthymine.

The product of Example 53B (4g; 10.7 mmoles), N-1-carboxymethylthymine (1.97 g; 10.7 mmoles). DhbtOH (1.75 g; 10.7 mmoles) and DIEA (1.9 ml; 11.2 mmoles) were dissolved in a mixture of $CH_2Cl_2$ (50ml) / DMF (50ml). After cooling at 0°C, DCC (2.2 g; 10.7 mmoles) was added and the reaction mixture was stirred at 0°C for 0.5h and 2.5h at room temperature. Then, DCU was filtered and washed with $CH_2Cl_2$ (2x100 ml). The combined organic phases were washed with 1M $NaHCO_3$ (3x100 ml), 1M $KHSO_4$ (3x100 ml), $H_2O$ (2x100 ml), dried over $Na_2SO_4$, filtered and evaporated under reduced pressure. The resulting oil was crystallized in chloroform/ether. (Yield : 4.04 g; 75%)

$^1$H-NMR ($CDCl_3$): δ (ppm): 1.4 (s,9H); 1.9 (s,3H); 3.4(m, 8H); 4.4 (d,2H); 5.1 (d,2H); 5.5 (m,1H); 5.9 (m,1H); 6.9 (s,1H): 7.4 (s,5H); 9.0 (s,1H). MS FAB+: M+1: 504

D.    Hydrogenolysis

The product of Example 53C (4 g; 7.9 mmoles) was dissolved in MeOH (150ml). At 0°C, 10% Pd/C (1.4 g) was added. The reaction mixture was hydrogenated at 0°C for 1h, filtered through Celite and the solvent removed under reduced pressure to give 4. (Yield: 3,0g; 100%)

$^1$H-NMR ($CDCl_3$): δ (ppm): 1.4 (s,9H); 1.9 (s,3H); 3.0-3.8(m); 4.6 (d,2H); 6.9 (s,1H); 7.3 (s,5H); 9.0 (s,1H). MS FAB+: M+1: 370.2

- 78 -

E. N-(benzyl acetate)-glycine ethyl ester hydrochloride.

At room temperature, benzyl bromoacetate (5.7 ml; 36 mmoles) was added over a period of 10 min to a solution of
5  glycine ethyl ester hydrochloride (5 g; 36 mmoles) and triethylamine (10.43 ml; 0.072 moles) dissolved in absolute ethanol (100 ml). After 4 days at room temperature, the solvent was removed under reduced pressure. The residue was dissolved in ethyl acetate (50 ml) / $H_2O$ (25 ml). After
10  separation of the two phases, the organic phase was thoroughly washed with water (8.25 ml). After evaporation of the solvent, the resulting oil was dissolved in ether (20 ml) / water (30 ml) and acidified to pH 4.5. After separation of the two phases, the aqueous phase was concentrated and the resulting
15  oil crystallized in cold ether. (Yield 4 g; 39%)

$^1$H-NMR (CDCl$_3$): δ (ppm): 1.3 (t,3H); 4.1 (d,4H); 4.4 (q,2H); 5.3 (d,2H); 5.5 (m,1H); 7.4 (s,5H). MS FAB+: M+1: 252.1

F. Coupling N-1-carboxymethylthymine.
20  The product of Example 53D (2.8g; 9.7 mmoles), N-1-carboxymethylthymine (1.79 g; 9.7 mmoles), DhbtOH (1.6g; 9.7 mmoles) and DIEA (1.7 ml; 10 mmoles) were dissolved in a mixture of $CH_2Cl_2$ (30 ml) / DMF (30 ml). After cooling at 0°C, DCC (2.0 g; 9.7 mmoles) was added and the reaction mixture was
25  stirred at 0°C for 0.5h and 4h at room temperature. Then DCU was filtered and wash with $CH_2Cl_2$ (120 ml). The combined organic phases were washed with 1M NaHCO$_3$ (3x60 ml), 1M KHSO$_4$ (3x60 ml), $H_2O$ (2x60 ml), dried over Na$_2$SO$_4$, filtered and evaporated under reduced pressure. The resulting oil was
30  crystallized in chloroform (60 ml) / petroleum ether (120 ml). (Yield: 3.235 g; 80%)

$^1$H-NMR (DMSO d6): δ (ppm): 1.3 (dt,3H); 1.9 (s,3H); 4.2 (m,4H); 4.4 (d,2H); 5.3 (d,2H); 7.4 (s,1H); 7.4 (s, 5H); 11.3 (s,1H). MS FAB+: M+1: 418.1

35  G. Hydrogenolysis
The product of Example 53E (4g; 9.6 mmoles) was dissolved in methanol (100ml) / DMF (25 ml). At 0°C, 10% Pd/C (1.6 g)

was added.  The reaction mixture was hydrogenated at 0°C for 1h, filtered through Celite and the solvent removed under reduced pressure.  (Yield = 3.12 g; 99%)

$^1$H-NMR (DMSO d6): δ (ppm): 1.3 (dt,3H); 1.9 (s,3H); 4.0-4.5 (m); 4.9 (s,2H); 7.2 (s,1H); 11.3 (s,1H).  MS FAB+: M+1: 328.1

### H.    Coupling

The products of Examples 53D (2.8g; 7.6 mmoles) and 53F (2.5g; 7.6 mmoles), and DhbtOH (1.24g; 7.6 mmoles) were dissolved in a mixture in a mixture of $CH_2Cl_2$ (50 ml) / DMF (50 ml).  After cooling at 0°C, DCC (1.56g; 7.6 mmoles) was added and the reaction mixture was stirred at 0°C for 0.5h and overnight at room temperature.  Then, DCU was filtered and washed with $CH_2Cl_2$ (250 ml).  The combined organic phases were washed with 1M $NaHCO_3$ (3x100 ml), 1M $KHSO_4$ (3x100 ml), $H_2O$ (2x100 ml), dried over $Na_2SO_4$, filtered and evaporated under reduced pressure.  The resulting oil was crystallized in chloroform (30 ml)/ petroleum ether (60 ml).  This material was recrystallized in chloroform (30 ml) / petroleum ether (25 ml) to give a pure compound.  (Yield: 0.60g; 12%)

MS FAB+: M+1: 418.1

### I.    Hydrolysis

The product of Example 53G (529 mg; 0.8 mmoles) was dissolved in absolute ethanol (100 ml).  Then NaOH 1M (20 ml) was added.  The reaction mixture was stirred for 30 min at room temperature and after cooling, neutralized with Dowex 50W (about 13g).  After filtration, washing with $H_2O$, ethanol and evaporation of the solvent, the residue was suspended in ether and filtered.  The acid product was obtained as a white solid, pure in HPLC (Yield: 450 mg: 86%)

MS FAB+: 673 (M+Na$^+$); 617 (M Na$^+$-tBu); 595 (M+1-tBu); 573 (M Na$^+$-Boc); 551 (M+1-Boc)

### EXAMPLE 54

### PNA Oligomer Having "Reversed" Monomer

The dimer prepared in Example 53 was incorporated into PNA H-TT(detT-idaT)CCTCTC-LysNH$_2$ (SEQ ID NO:53) generally in accordance with Example 1.  The melting temperatures, $T_m$, of

complexes between this PNA and decamers 5'd(AAAAGGAGAG) (SEQ ID
NO:54) and 5'd(GAGAGGAAAA) (SEQ ID NO:55) were 55°C and 43.5°C,
respectively, at pH 7.  By comparision, the $T_m$ of complexes
between  PNA  H-TTTCCTCTC-LysNH$_2$  (SEQ  ID  NO:56),  prepared

5  generally in accordance with Example 1, and these decamers was
58.5°C and 40.5°C, respectively, at pH 7.

Those skilled in the art will appreciate that numerous
changes  and  modifications  may  be  made  to  the  preferred
embodiments  of  the  invention  and  that  such  changes  and

10  modifications may be made without departing from the spirit of
the  invention.   It  is  therefore  intended  that  the  appended
claims cover all such equivalent variations as fall within the
true spirit and scope of the invention.

- 81 -

WE CLAIM:

1.    A complex comprising a first nucleic acid strand and second and third strands;

said second and third strands independently containing a sequence of ligands covalently bound by linking moieties;

at least one of said second strand linking moieties comprising an amide, a thioamide, a sulfinamide or a sulfonamide linkage and a plurality of said ligands on said second strand interacting with said first strand; and

at least one of said third strand linking moieties comprising an amide, a thioamide, a sulfinamide or a sulfonamide linkage and a plurality of said ligands on said third strand interacting with said first strand or with said ligands on said second strand.

2.    The complex of claim 1 wherein each of said second strand and said third strand, independently, include a plurality of monomeric units connected via amide, thioamide, sulfinamide or sulfonamide linkages.

3.    The complex of claim 1 wherein said first strand is a DNA strand.

4.    The complex of claim 1 wherein said first stand is a RNA strand.

5.    The complex of claim 2 wherein said monomeric units are connected via amide linkages.

6.    The complex of claim 2 wherein said monomeric units each include a first nitrogen that connects said ligands to said monomeric units and a second nitrogen atom that forms a portion of said amide, thioamide, sulfinamide or sulfonamide linkage.
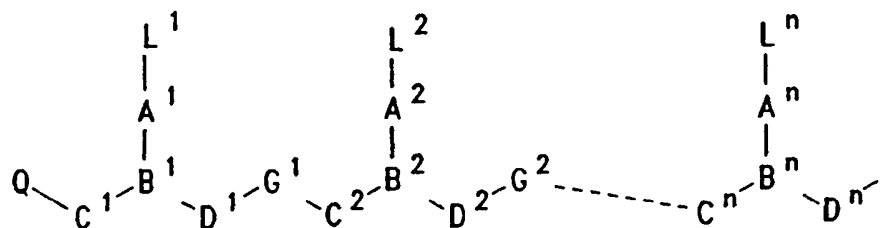
7.    The complex of claim 2 wherein a majority of said ligands of at least one of said second strand or third strand are pyrimidine bases.

8.    The complex of claim 7 wherein said pyrimidine bases are selected from thymine and cytosine.

9.    The complex of claim 2 wherein said nucleic acid strand is a purine rich strand.

10.    The complex of claim 2 wherein said monomeric units comprise aminoethylglycine units.

11.    The complex of claim 1 wherein said second and said third strands, independently, comprise a compound of the formula:



wherein:

n is at least 2,

each of $L^1$-$L^n$ is independently selected from the group consisting of hydrogen, hydroxy, $(C_1$-$C_4)$alkanoyl, naturally occurring nucleobases, non-naturally occurring nucleobases, aromatic moieties, DNA intercalators, nucleobase-binding groups, heterocyclic moieties, and reporter ligands;

each of $C^1$-$C^n$ is $(CR^6R^7)_y$ where $R^6$ is hydrogen and $R^7$ is selected from the group consisting of the side chains of naturally occurring alpha amino acids, or $R^6$ and $R^7$ are independently selected from the group consisting of hydrogen, $(C_2$-$C_6)$alkyl, aryl, aralkyl, heteroaryl, hydroxy, $(C_1$-$C_6)$alkoxy, $(C_1$-$C_6)$alkylthio, $NR^3R^4$ and $SR^5$, where $R^3$ and $R^4$ are as defined above, and $R^5$ is hydrogen, $(C_1$-$C_6)$alkyl, hydroxy-, alkoxy-, or alkylthio- substituted $(C_1$-$C_6)$alkyl, or $R^6$ and $R^7$ taken together complete an alicyclic or heterocyclic system;

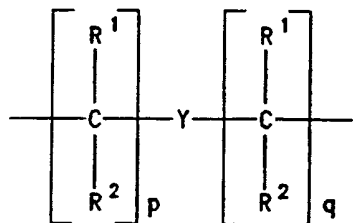each of $D^1$-$D^n$ is $(CR^6R^7)_z$ where $R^6$ and $R^7$ are as defined above;

each of y and z is zero or an integer from 1 to 10, the sum y + z being greater than 2 but not more than 10;

each of $G^1$-$G^{n-1}$ is -$NR^3CO$-, -$NR^3CS$-, -$NR^3SO$- or -$NR^3SO_2$-, in either orientation, where $R^3$ is as defined above;
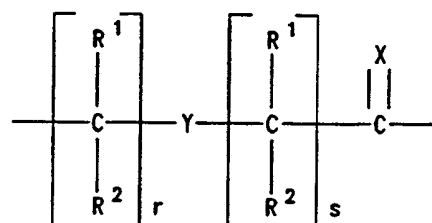
each of $A^1$-$A^n$ and $B^1$-$B^n$ are selected such that:

(a)   A is a group of formula (IIa), (IIb), (IIc) or (IId), and B is N or $R^3N^+$; or
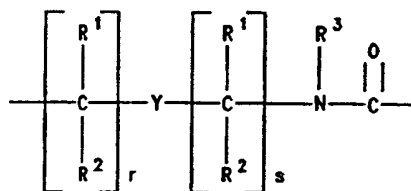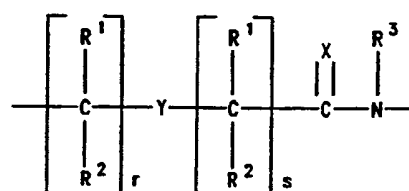
(b)   A is a group of formula (IId) and B is CH;

$$\left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array}\right]_p \ -Y- \ \left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array}\right]_q$$

(IIa)

$$\left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array}\right]_r \ -Y- \ \left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array}\right]_s \ \begin{array}{c} X \\ || \\ -C- \end{array}$$

(IIb)

$$\left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array}\right]_r \ -Y- \ \left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array}\right]_s \ \begin{array}{c} R^3 \ \ O \\ | \ \ \ || \\ -N-C- \end{array}$$

(IIc)

$$\left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array}\right]_r \ -Y- \ \left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array}\right]_s \ \begin{array}{c} X \ \ R^3 \\ || \ \ \ | \\ -C-N- \end{array}$$

(IId)

where:

X is O, S, Se, $NR^3$, $CH_2$ or $C(CH_3)_2$;

Y is a single bond, O, S or $NR^4$;

each of p and q is zero or an integer from 1 to 5, the sum p+q being not more than 10;

each of r and s is zero or an integer from 1 to 5, the sum r+s being not more than 10;

each $R^1$ and $R^2$ is independently selected from the group consisting of hydrogen, $(C_1$-$C_4)$alkyl which may be hydroxy- or alkoxy- or alkylthio-substituted, hydroxy, alkoxy, alkylthio, amino and halogen; and
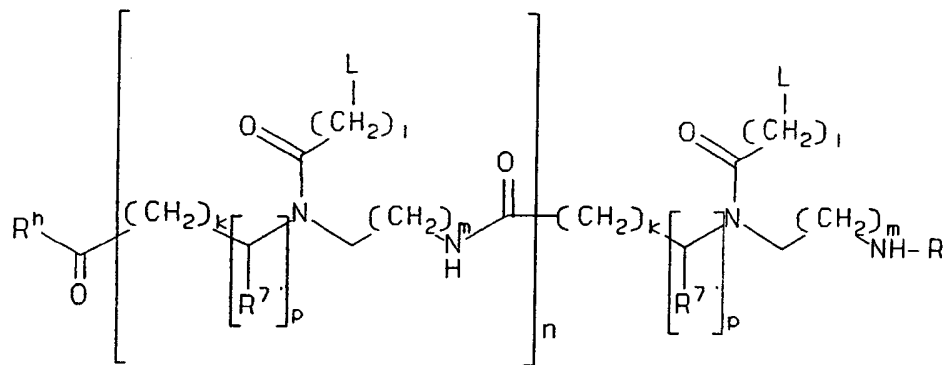
each $R^3$ and $R^4$ is independently selected from the group consisting of hydrogen, $(C_1$-$C_4)$alkyl, hydroxy- or alkoxy- or alkylthio-substituted $(C_1$-$C_4)$alkyl, hydroxy, alkoxy, alkylthio and amino;

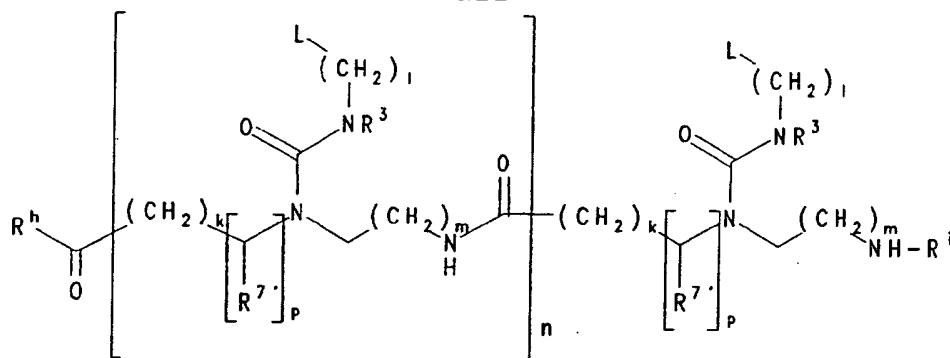Q is $-CO_2H$, $-CONR'R''$, $-SO_3H$ or $-SO_2NR'R''$ or an activated derivative of $-CO_2H$ or $-SO_3H$; and

I is $-NHR'''R''''$ or $-NR'''C(O)R''''$, where R', R'', R''' and R'''' are independently selected from the group consisting of hydrogen, alkyl, amino protecting groups,

- 84 -

reporter ligands, intercalators, chelators, peptides, proteins,
carbohydrates, lipids, steroids, nucleosides, nucleotides,
nucleotide diphosphates, nucleotide triphosphates, oligonucleo-
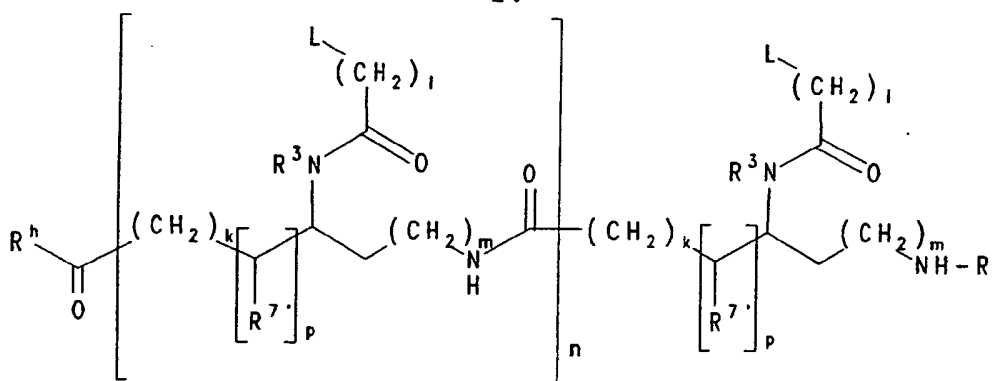tides, oligonucleosides and soluble and non-soluble polymers.

    12.   The complex of claim 1 wherein said second strand
and said third strand, independently, comprise a compound of
the formula III, IV or V:
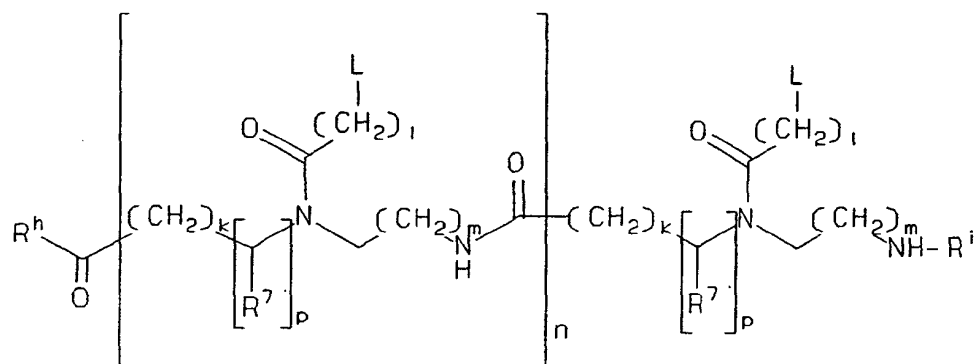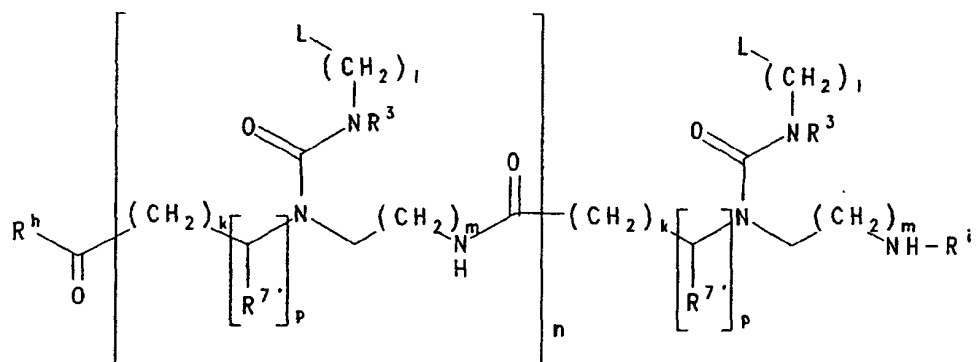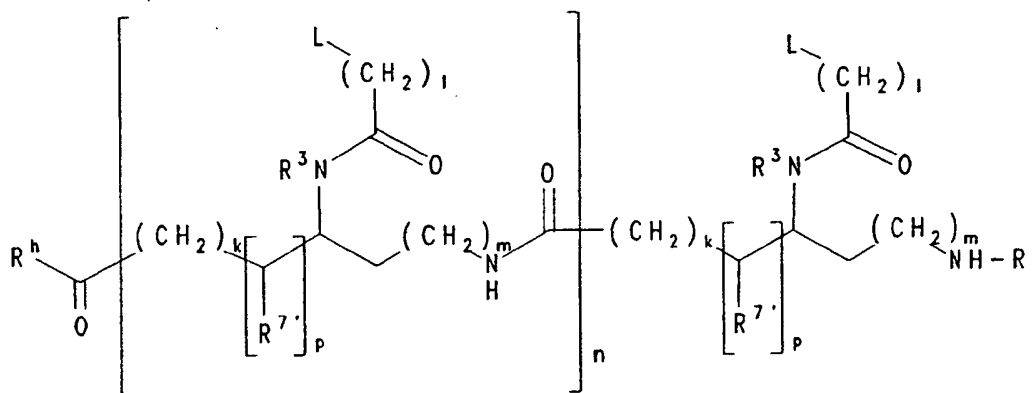


III



IV

V

wherein:

each L is independently selected from the group consisting of hydrogen, phenyl, heterocyclic moieties, naturally occurring nucleobases, and non-naturally occurring nucleobases;

each $R^{7'}$ is independently selected from the group consisting of hydrogen and the side chains of naturally occurring alpha amino acids;

n is an integer greater than 1, each k, l, and m is, independently, zero or an integer from 1 to 5;

each p is zero or 1;

$R^h$ is OH, $NH_2$ or $-NHLysNH_2$; and

$R^i$ is H or $COCH_3$.

13. The complex of claim 1 wherein said interaction between said first and second strands occurs between heteocyclic bases of said first strand and ligands on said second strand.

14. The complex of claim 1 wherein said interaction between said first and third strands occurs between heteocyclic bases of said first strand and ligands on said third strand.

15. The complex of claim 1 wherein said interaction includes formation of non-covalent bonds.

16. The complex of claim 1 wherein said interaction of ligands is via hydrogen bonding.

17. The complex of claim 1 wherein said interaction between said first and second strands occurs between

heteocyclic bases of said first strand and ligands on said second strand;

said interaction between said first and third strands occurs between heteocyclic bases of said first strand and ligands on said third strand; and

said interaction of ligands of said second strand and heterocyclic bases of said first strand is via Watson/Crick hydrogen bonding, and said interaction of ligands of said second strand with heterocyclic bases of said first strand is via one of Watson/Crick hydrogen bonding or Hoogsteen hydrogen bonding.

18. A complex comprising a first nucleic acid strand, a second nucleic acid strand, and a third strand;

said third strand containing a sequence of ligands covalently bound by linking moieties and wherein a plurality of said ligands interact with at least one of said first strand and said second strand; and

at least one of said third strand linking moieties comprises an amide, a thioamide, a sulfinamide or asulfonamide linkage.

19. The complex of claim 18 wherein said third strand includes a plurality of monomeric units that are connected via amide, thioamide, sulfinamide or sulfonamide backbone linkages.

20. The complex of claim 18 wherein said interaction includes formation of non-covalent bonds.

21. The complex of claim 20 wherein said interaction includes hydrogen bonding.

22. The complex of claim 18 wherein said interaction occurs between heteocyclic bases on said first strand or said second strand and ligands on said third strand.

23. The complex of claim 18 wherein said first strand and said second strand are DNA strands.

24. The complex of claim 23 wherein said DNA strands are double-stranded DNA.

25. The complex of claim 23 wherein said DNA strands are complementary DNA strands.

26.   The complex of claim 19 wherein said monomeric units are connected via amide linkages.

27.   The complex of claim 19 wherein said monomeric units each include a first nitrogen that connects said ligands to said monomeric units and a second nitrogen atom that forms a portion of said amide, thioamide, sulfinamide or sulfonamide linkage.

28.   The complex of claim 18 wherein a majority of said ligands are pyrimidine bases.

29.   The complex of claim 28 wherein said pyrimidine bases are selected from thymine and cytosine.

30.   The complex of claim 18 wherein one of said nucleic acid strands is a purine rich strand.

31.   The complex of claim 19 wherein said monomeric units comprise aminoethylglycine units.

32.   The complex of claim 18 wherein said third strand comprises a compound of the formula:



wherein:

n is at least 2,

each of $L^1$-$L^n$ is independently selected from the group consisting of hydrogen, hydroxy, $(C_1$-$C_4)$alkanoyl, naturally occurring nucleobases, non-naturally occurring nucleobases, aromatic moieties, DNA intercalators, nucleobase-binding groups, heterocyclic moieties, and reporter ligands;

each of $C^1$-$C^n$ is $(CR^6R^7)_y$ where $R^6$ is hydrogen and $R^7$ is selected from the group consisting of the side chains of naturally occurring alpha amino acids, or $R^6$ and $R^7$ are independently selected from the group consisting of hydrogen, $(C_2$-$C_6)$alkyl, aryl, aralkyl, heteroaryl, hydroxy, $(C_1$-$C_6)$alkoxy, $(C_1$-$C_6)$alkylthio, $NR^3R^4$ and $SR^5$, where $R^3$ and $R^4$ are as defined above, and $R^5$ is hydrogen, $(C_1$-$C_6)$alkyl, hydroxy-, alkoxy-, or

alkylthio- substituted $(C_1-C_6)$alkyl, or $R^6$ and $R^7$ taken together complete an alicyclic or heterocyclic system;

each of $D^1-D^n$ is $(CR^6R^7)_z$ where $R^6$ and $R^7$ are as defined above;

each of y and z is zero or an integer from 1 to 10, the sum y + z being greater than 2 but not more than 10;

each of $G^1-G^{n-1}$ is $-NR^3CO-$, $-NR^3CS-$, $-NR^3SO-$ or $-NR^3SO_2-$, in either orientation, where $R^3$ is as defined above;

each of $A^1-A^n$ and $B^1-B^n$ are selected such that:

(a)  A is a group of formula (IIa), (IIb), (IIc) or (IId), and B is N or $R^3N^+$; or

(b)  A is a group of formula (IId) and B is CH;

(IIa)          (IIb)

(IIc)          (IId)

where:

X is O, S, Se, $NR^3$, $CH_2$ or $C(CH_3)_2$;

Y is a single bond, O, S or $NR^4$;

each of p and q is zero or an integer from 1 to 5, the sum p+q being not more than 10;

each of r and s is zero or an integer from 1 to 5, the sum r+s being not more than 10;

each $R^1$ and $R^2$ is independently selected from the group consisting of hydrogen, $(C_1-C_4)$alkyl which may be hydroxy- or alkoxy- or alkylthio-substituted, hydroxy, alkoxy, alkylthio, amino and halogen; and

each $R^3$ and $R^4$ is independently selected from the group consisting of hydrogen, $(C_1-C_4)$alkyl, hydroxy- or alkoxy- or alkylthio-substituted $(C_1-C_4)$alkyl, hydroxy, alkoxy, alkylthio and amino;

Q is $-CO_2H$, $-CONR'R''$, $-SO_3H$ or $-SO_2NR'R''$ or an activated derivative of $-CO_2H$ or $-SO_3H$; and

I is $-NHR'''R''''$ or $-NR'''C(O)R''''$, where $R'$, $R''$, $R'''$ and $R''''$ are independently selected from the group consisting of hydrogen, alkyl, amino protecting groups, reporter ligands, intercalators, chelators, peptides, proteins, carbohydrates, lipids, steroids, nucleosides, nucleotides, nucleotide diphosphates, nucleotide triphosphates, oligonucleotides, oligonucleosides and soluble and non-soluble polymers.

- 90 -

33.  The complex of claim 18 wherein said third strand comprises a compound of the formula III, IV or V:



III



IV



V

wherein:

each L is independently selected from the group consisting of hydrogen, phenyl, heterocyclic moieties, naturally occurring nucleobases, and non-naturally occurring nucleobases;

each $R^{7'}$ is independently selected from the group consisting of hydrogen and the side chains of naturally occurring alpha amino acids;

n is an integer greater than 1,

each k, l, and m is, independently, zero or an integer from 1 to 5;8

each p is zero or 1;

$R^h$ is OH, $NH_2$ or $-NHLysNH_2$; and

$R^i$ is H or $COCH_3$.

34. A process for modifying double-stranded DNA, comprising the steps of:

contacting said double-stranded DNA with a compound that comprises a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, wherein a plurality of said ligands interact with said double-stranded DNA and thereby displace one of said DNA strands; and modifying said displaced strand.

35. The process of claim 34 wherein modifying said displaced strand comprises cleaving said strand.

36. The process of claim 35 wherein said cleavage is effected by an enzyme.

37. The process of claim 36 wherein said double-stranded DNA is contacted with said compound intercellular and said enzymatic cleavage is effect by enzymes naturally present intercellular.

38. The process of claim 37 wherein said enzyme is a nuclease.

39. The process of claim 38 wherein said nuclease is nuclease $S_1$.

40. The process of claim 34 wherein said compound further includes a moiety that modifies said displaced strand.

41. The process of claim 40 wherein moiety cleaves said displaced strand.

42. The process of claim 34 wherein a plurality of said ligands hydrogen bond with said DNA.

43.  The process of claim 34 wherein said compound forms Watson-Crick hydrogen bonds with one strand of said double-stranded DNA.

44.  The process of claim 31 wherein a first molecule of said compound forms Watson-Crick hydrogen bonds with a first strand of said double-stranded DNA and a second molecule of said compound forms Hoogsteen hydrogen bonds with said first strand.

45.  The process of claim 34 wherein modifying said displaced strand comprises covalently bonding a bonding moiety to said displaced strand.

46.  The process of claim 34 wherein modifying said displaced strand comprises activating cellular repair mechanisms to effects digestion of said displaced strand.

47.  The process of claim 34 wherein said displaced strand is modified by covalently bonding a bonding moiety to said displaced stand, thereby initiating a cellular repair mechanism that effects digestion of said displaced strand.

48.  A process of inhibiting the expression of a gene comprising contacting said gene with a compound that comprises a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, and wherein a plurality of said ligands interact with said gene.

49.  The process of claim 48 wherein a plurality of said ligands hydrogen bond with DNA of said gene.

50.  The process of claim 48 wherein inhibition of gene expression comprises transcription interference.

51.  The process of claim 50 wherein said transcription interferences comprises RNA polymerase arrest.

52.  The process of claim 51 wherein said transcription interference comprises transcription arrest outside of a promoter region of said gene.

53.  The process of claim 49 wherein a plurality of said ligands hydrogen bond to a template strand of said gene.

54.  The process of claim 48 wherein said compound includes at least six monomeric units connected via a polyamide

- 93 -

backbone and each of said monomeric units includes one of said
ligands.

56. A process for arresting transcription of a gene
comprising contacting said gene with a compound that comprises
a sequence of ligands covalently bound by amide, thioamide,
sulfinamide, or sulfonamide linking moieties, thereby binding
a plurality of said ligands to a template DNA strand of said
gene.

56. A process for modulating the activity of a
restriction enzyme at a DNA restriction site, comprising
contacting said DNA with a compound that comprises a sequence
of ligands covalently bound by amide, thioamide, sulfinamide,
or sulfonamide linking moieties, thereby binding a plurality of
said ligands to said DNA proximal to said restriction site.

57. A process for sequencing DNA, comprising:
contacting said DNA with a compound that comprises a sequence
of ligands covalently bound by amide, thioamide, sulfinamide,
or sulfonamide linking moieties, thereby binding a plurality of
said ligands to said DNA proximal to a DNA restriction site;
treating said DNA with a restriction enzyme that recognizes and
cleaves said DNA at said restriction site; and
identifying at least one product of said cleavage.

58. A process for inhibiting transcription of DNA
comprising:
selecting DNA for which transcription is to be
inhibited; and
adding to a mixture containing said DNA and a transcription
factor for said DNA a compound that comprises a sequence of
ligands covalently bound by amide, thioamide, sulfinamide, or
sulfonamide linking moieties, thereby binding a plurality of
said ligands to said transcription factor.

59. A method for modulating binding of RNA polymerase
to double-stranded DNA comprising:
contacting said DNA with a compound that comprises a sequence
of ligands covalently bound by amide, thioamide, sulfinamide,
or sulfonamide linking moieties, thereby binding a plurality of
said ligands to said DNA; and exposing said complex to RNA polymerase.

60.   A process for initiating transcription of a gene comprising contacting said gene with a compound that comprises a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, wherein a plurality of said ligands interact with said gene to melt double-stranded DNA of said gene and thereby expose a template strand of said DNA to recognition by RNA polymerase.

61.   The process of claim 60 wherein said gene is contacted with said compound under conditions that form a hybrid strand with a non-coding strand of said DNA.

62.   The process of claim 60 wherein said gene is contacted with said compound under conditions that form both a hybrid strand with said DNA and a template strand "D" loop.

63.   A process for initiating transcription of a gene, comprising contacting said gene with a compound that comprises a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, wherein a plurality of said ligands interact with said gene to melt double-stranded DNA of said gene and thereby form a transcription elongation loop.

64.   A method for binding RNA polymerase to double-stranded DNA comprising:

contacting said DNA with a compound that comprises a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, wherein a plurality of said ligands interact with said DNA; and

exposing said compound and said DNA to said RNA polymerase.

65.   A hybrid complex for modulating transcription comprising:

double-stranded DNA;

a compound that comprises a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, wherein a plurality of said ligands interact with said double-stranded DNA; and

RNA polymerase in contact with at least one of said DNA and said compound.

66. A synthetic transcription factor comprising: double-stranded DNA; and

a compound that comprises a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, wherein a plurality of said ligands interact with said double-stranded DNA.

67. A sequence specific gene activator comprising a compound that comprises a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, wherein a plurality of said ligands interact with a specific sequence of DNA of said gene.

68. A sequence specific gene activator comprising first and second synthetic strands, wherein each strand comprises a sequence of about 6-50 ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, and each of said strands has a ligand sequence that recognizes a base sequence within proximal regions on at least one DNA strand of said gene such that said strands bind proximal to one another on said DNA strand.

69. The gene activator of claim 68 wherein the ligand sequence of said synthetic strands are complementary to a non-template strand of said DNA.

70. The gene activator of claim 68 wherein the ligand sequence of said synthetic strands is selected such that binding of said synthetic strands on said gene results in creation of a loop of DNA that includes the binding sites of said synthetic strands on said gene.

71. A sequence specific gene activator comprising first and second synthetic strands, wherein:

each strand comprises a sequence of about 6-50 ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties;

said first strand has a ligand sequence that recognizes a base sequence on one of the DNA strands of said gene;

- 96 -

said second strand has a ligand sequence that recognizes a base sequence on the other of the DNA strands of said gene; and

the recognized base sequence on said one DNA strand and the recognized base sequence of said other DNA strand are positioned proximal to one another on said gene.

72.  The gene activator of claim 71 wherein ligand sequences of said synthetic strands is selected such that binding of said synthetic strands on said gene results in creation of a loop of DNA on said gene that includes the binding sites of said synthetic strands on said gene.

73.  A chimeric compound comprising a first strand section including DNA or RNA, and a second strand section including a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, said first strand section and said second strand section being covalently bound.

74.  The chimeric structure of claim 73 wherein said first stand section includes DNA.

75.  A double-stranded chimeric complex, wherein:

a first strand comprises a first strand section including DNA or RNA and a second strand section comprises a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, said first and said second strand sections being covalently bound;

a second strand comprises DNA, RNA, or a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties; and

said first strand and said second strand interact with one another to form a duplex structure.

76.  A method for intercellularly modulating the activity of a transcription factor comprising:

forming a chimeric structure by covalently bonding a first strand section including DNA or RNA and a second strand section including a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties,

wherein a plurality of said ligands bind with said transcription factor; and

introducing said chimeric structure into a cell.

77. The method of claim 76 further including selecting said ligands and said DNA or RNA such that a plurality of said ligands bind with the sugar-phosphate backbone of said RNA or DNA.

78. A process for intercellularly inhibiting the binding of a protein comprising:

forming a compound that comprises a sequence of ligands covalently bound by amide, thioamide, sulfinamide, or sulfonamide linking moieties, wherein a plurality of said bind to said protein; and

introducing said structure into a cell.

FIG. 1
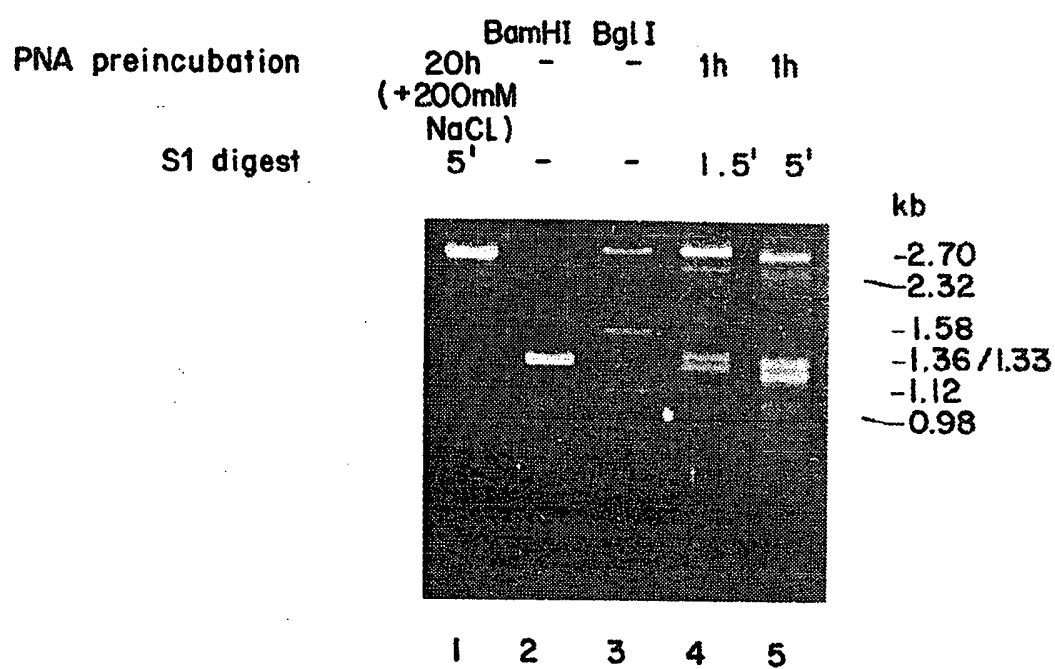
| DNA: | pUC19 | | pT9C | | | | | | | pT9C-5 |
|---|---|---|---|---|---|---|---|---|---|---|
| PNA: | – | 1 | 1 | 1 | – | 1 | – | 2 | 3 | 1 |
| Enzyme: | BamH1 | | | | S1 | | | | | |
|  |  | 15 | 30 | 30 | 0 | 0 | 30 | 30 | 30 | 30 |
|  |  |  | | +Na | | | | | | |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
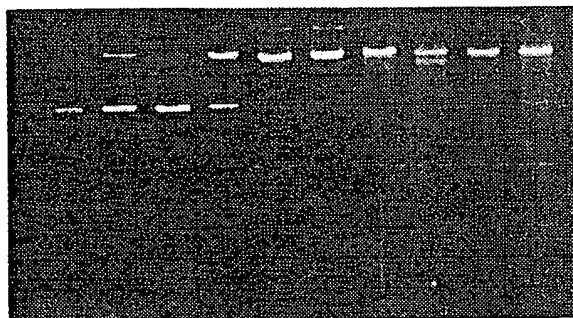


FIG. 2
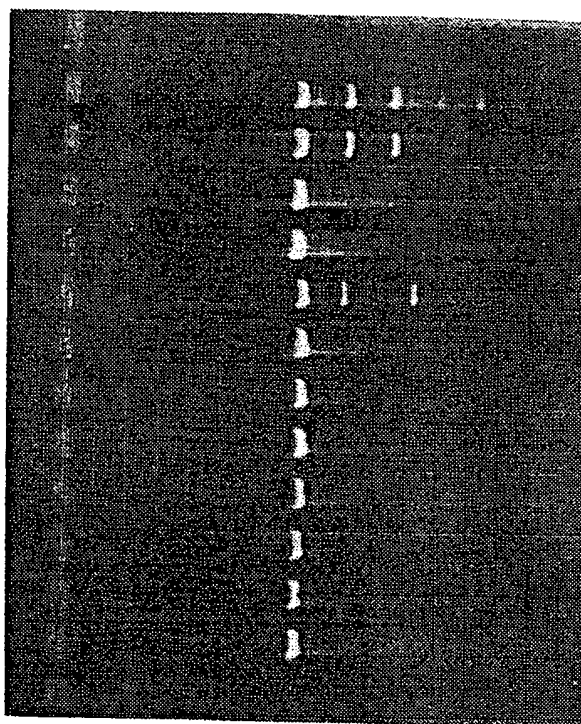
FIG. 3
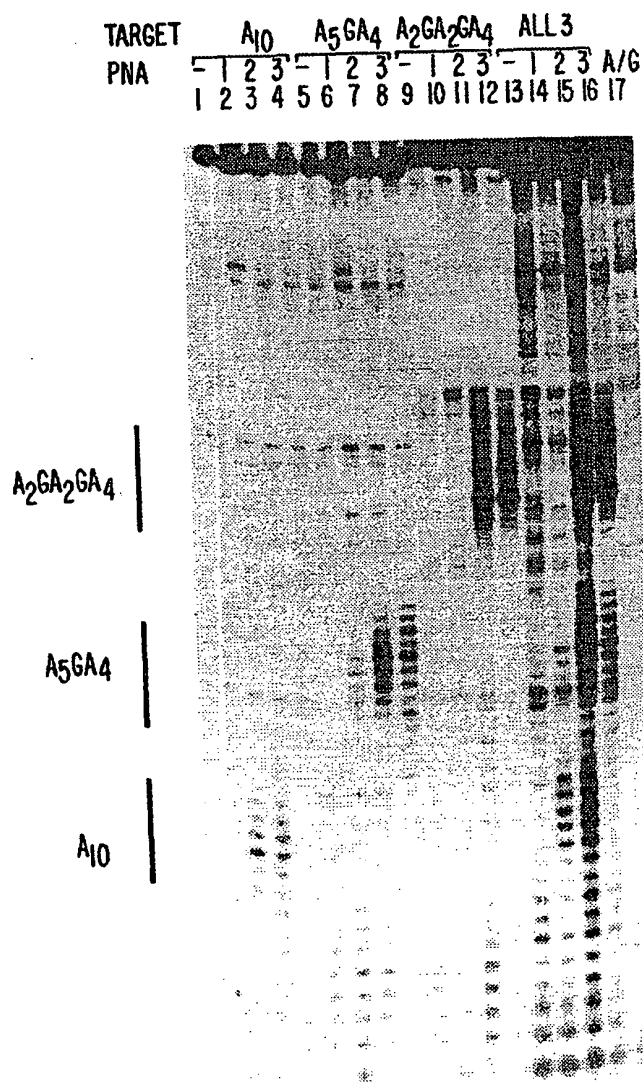
PNA 1: $T_{10}$     PNA 2: $T_5CT_4$     PNA 3: $T_2CT_2CT_4$

FIG. 4

PNA/DNA    0   0.006   0.02   0.06   0.2   0.6
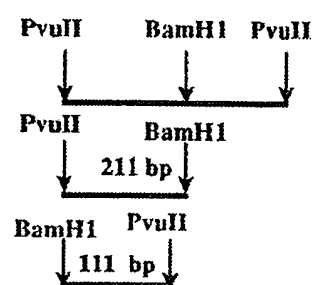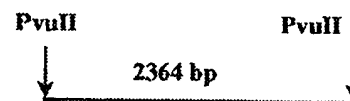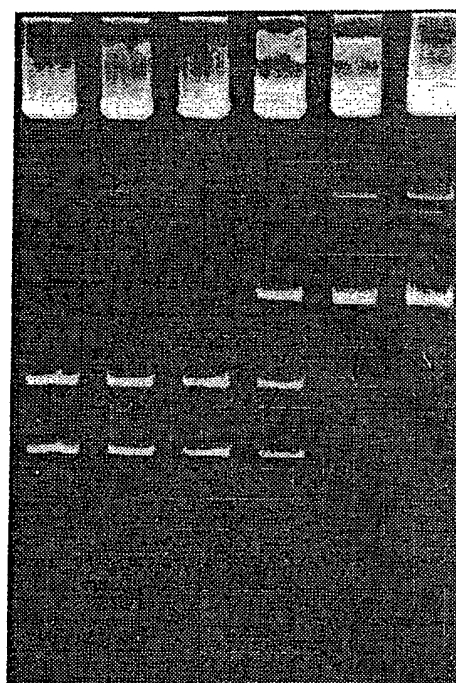


PNA Target

5´------GGATCCAAAAAAAAAAGGATCC--------
3´------CCTAGGTTTTTTTTTTCCTAGG--------

BamH1           BamH1

FIG.5

FIG. 6



$T_{10}$

$T_8$

$T_6$

Dissociation Temperature
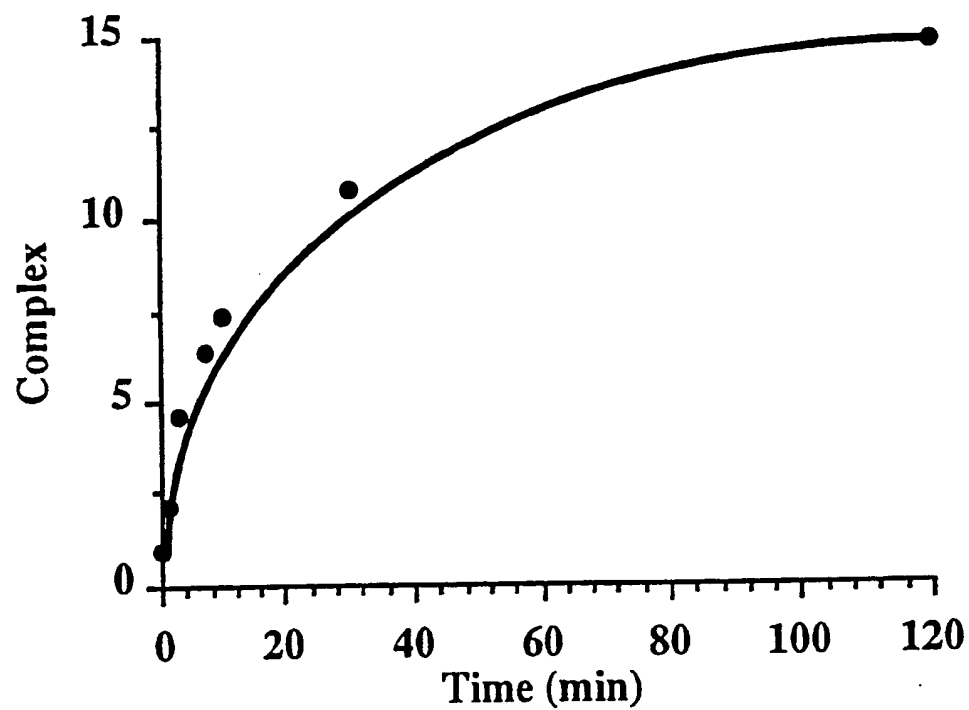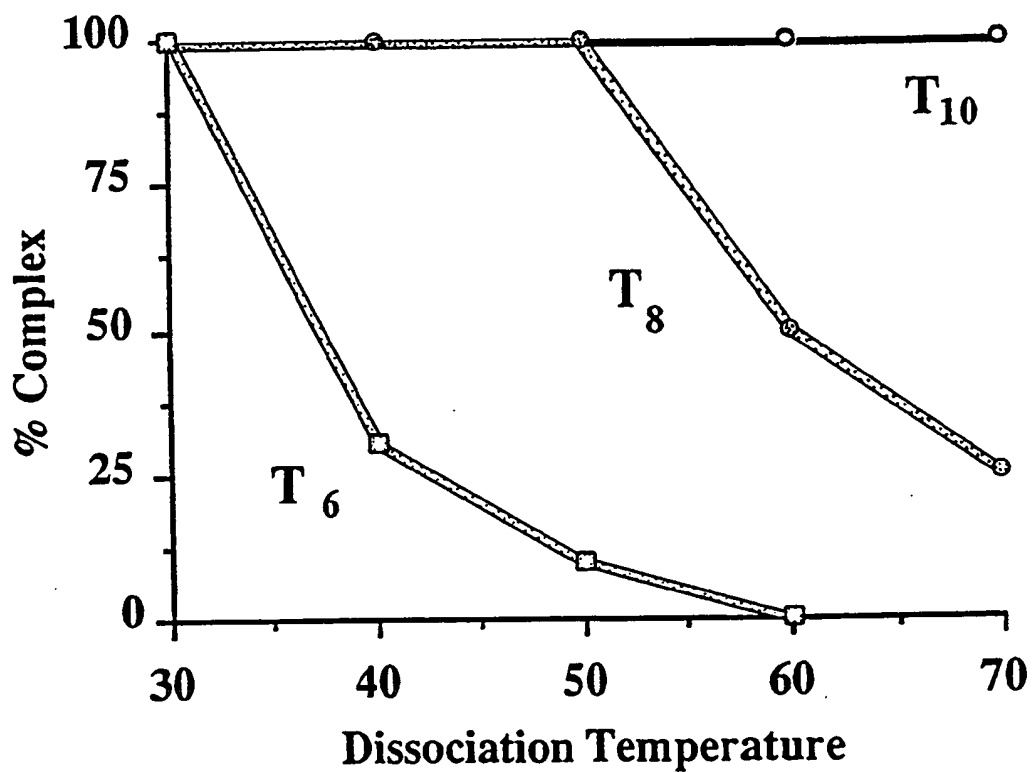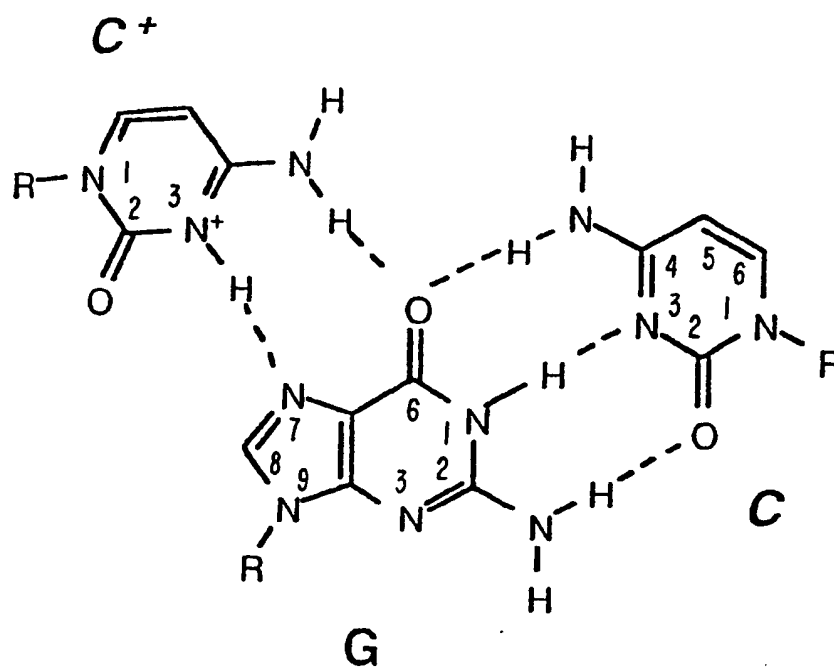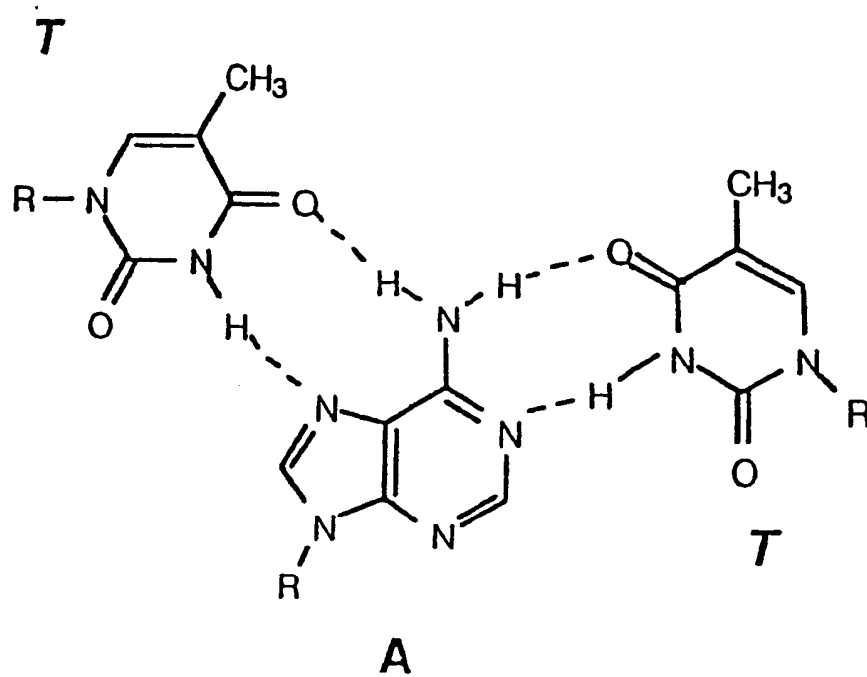
FIG. 7          SUBSTITUTE SHEET (RULE 26)

FIG. 8
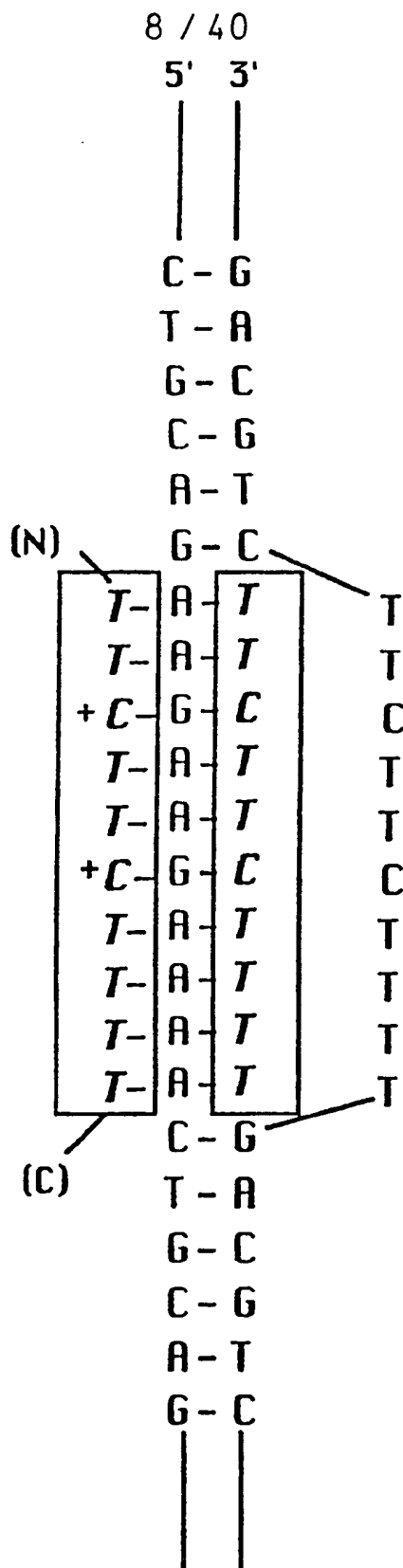
FIG. 9
SUBSTITUTE SHEET (RULE 26)

9 / 40
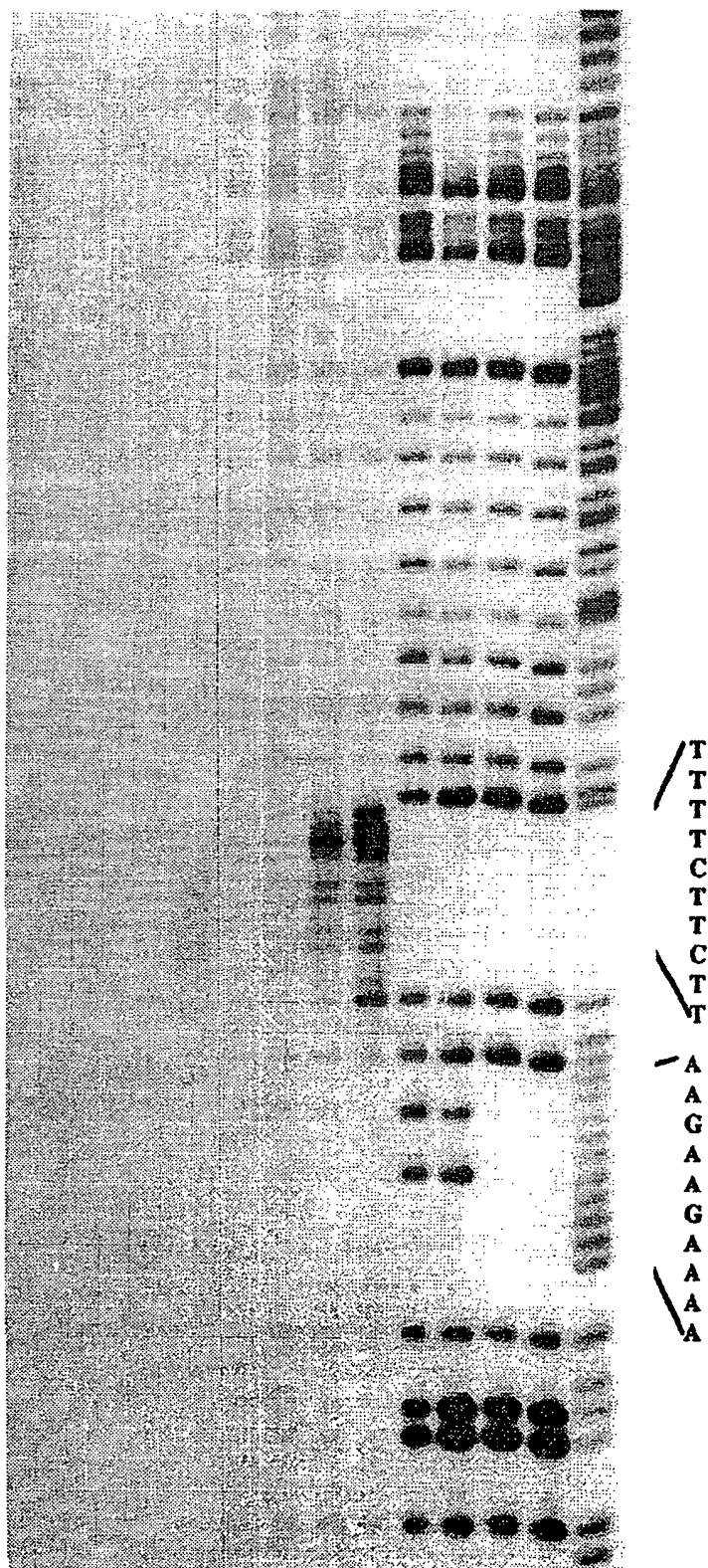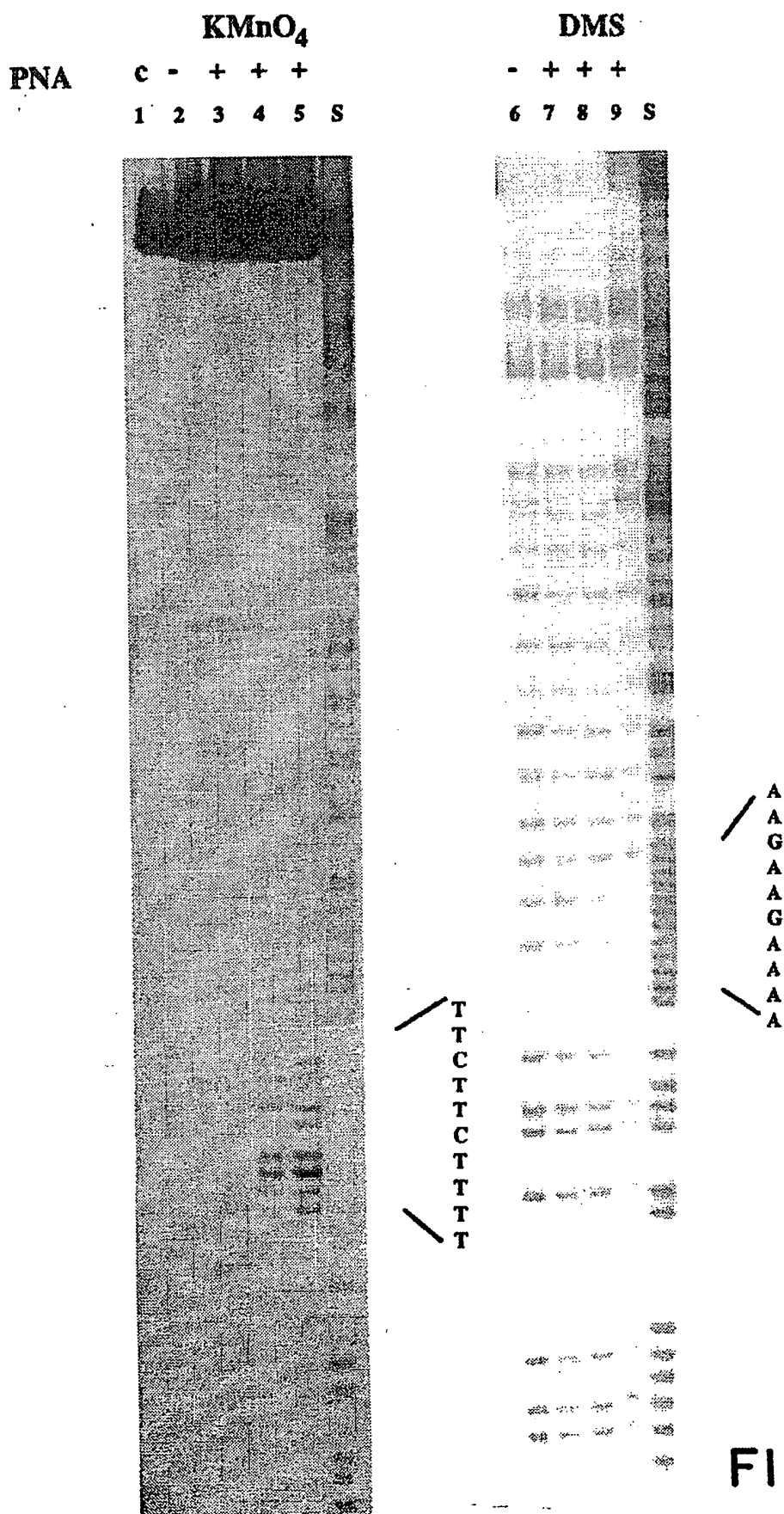


FIG.10

FIG.H

FIG.12

FIG.13A          FIG.13B



FIG.13C

FIG. 14

FIG. 15



1st Dimension

2nd Dimension (Chloroquine)

FIG. 16

FIG. 17A



FIG. 17B

FIG. 18

FIG.19

FIG. 20

FIG.21

FIG. 22

FIG. 23

FIG.24

PNA   - - - + +    - - - + +    - - - + +
      1 2 3 4 5    1 2 3 4 5    1 2 3 4 5

PvuII

XbaI

BamHI

# FIG.25a FIG.25b FIG.25c

FIG. 26

FIG. 27

FIG. 28

FIG. 29

lacUV5

PvuII  lacUV5

PvuII

lacUV5-RNA

PNA-RNA

FIG. 30

pCH495

| CMV | | TTTTCTTTT | |
AAAAGAAAA
373    495
HIV

pCIN

| CMV | | CCCCCCCTTTTTTTTTT | |
GGGGGGGAAAAAAAAAA
272    481
CMV InB

pCINr.

| CMV | | AAAAAAAAAGGGGGGG | |
TTTTTTTTTCCCCCCC
209    481
CMV InB (rev.)

pIE72

| CMV | | CCCTTCCTCC | |
GGGAAGGAGG
206    318

pIE72

| CMV | | | TCTCTCTCTT | |
AGAGAGAGAA
1100    1200
CMV cDNA

FIG. 31

SUBSTITUTE SHEET (RULE 26)

TTTTCTTTT          Control

| [PNA uM] | 0 | 0 | 1000 | 100 | 10 | 1 | 1000 | 100 | 10 | 1 |

495 ➡
373 ➡

190 ➡

## FIG. 32

FIG. 33

[KCl mM]    100    50    10    5    1    0.5

495 ➡

373 ➡

% truncated

| 0 | 4 | 28 | 57 | 65 | 71 |
|---|---|----|----|----|----|

# FIG.34

FIG.35

FIG.36B



FIG. 36A
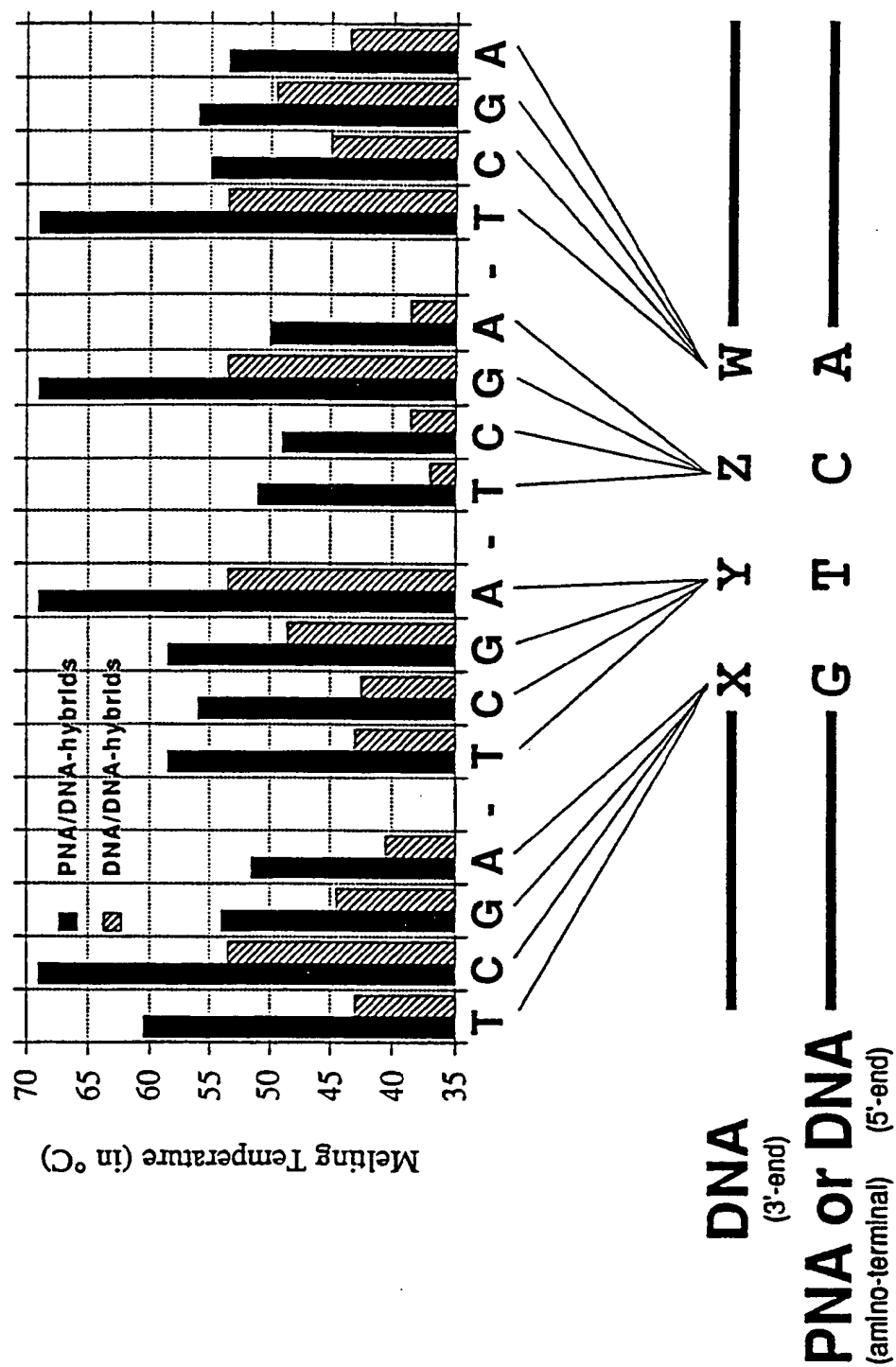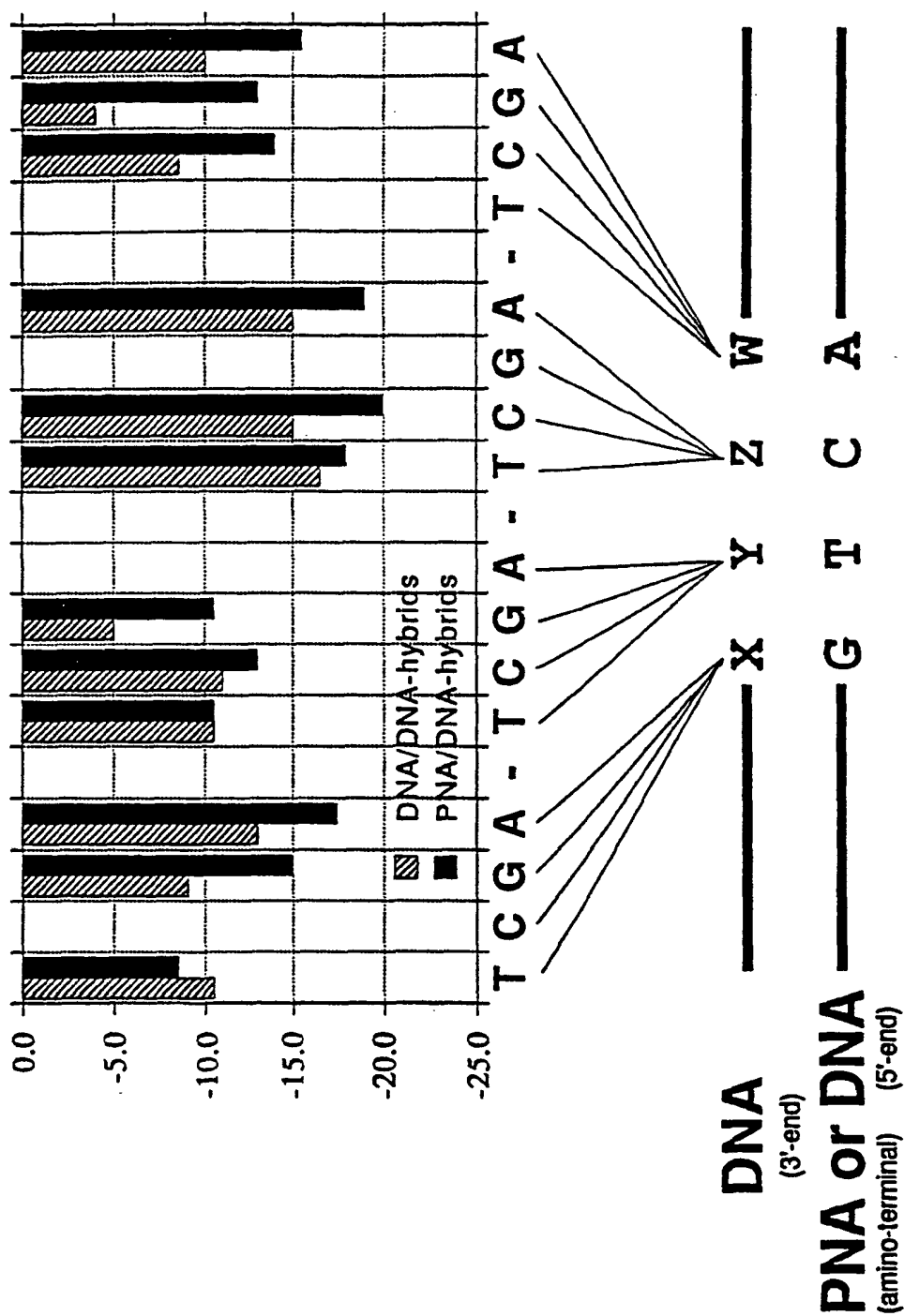
FIG. 37

FIG. 38

PNA-DNA
COMPLEX

DNA
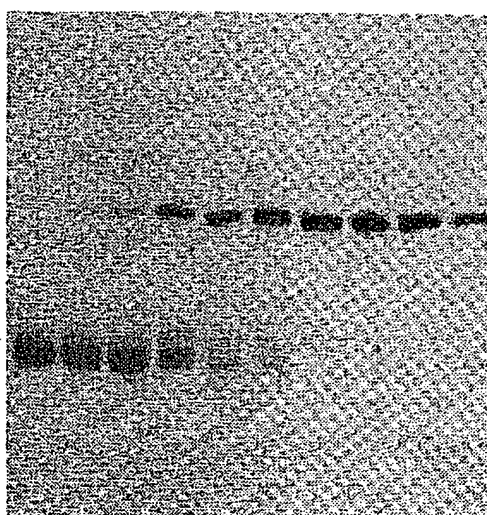
0  0.25 0.33 0.5 0.7  I  1.25 1.5 2.5 5
PNA:DNA RATIO

# FIG. 39

FIG.40

a: DNA/DNA 5'-TGA ACG TCA CAA CTA-3'
3'-ACA TGC AGT GTT GAT-5'

b: PNA/DNA (antiparallel)

c: PNA/DNA (parallel)
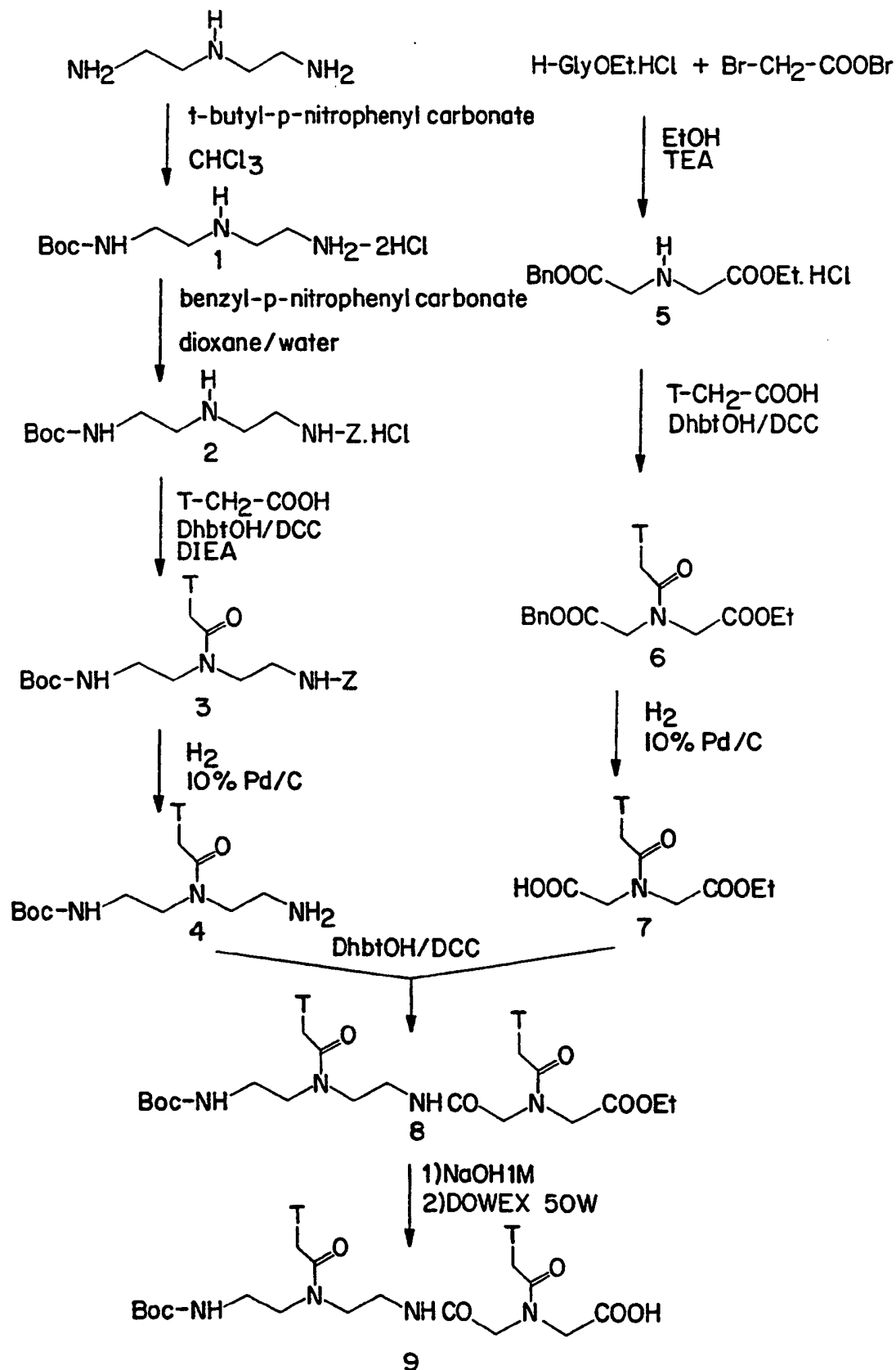
FIG. 41

d: DNA/RNA 5'-TGT ACG TCA CAA CTA-3'
3'-ACA UGC AGU GUU GAU-5'

e: PNA/RNA (antiparallel)

f: PNA/RNA (parallel)
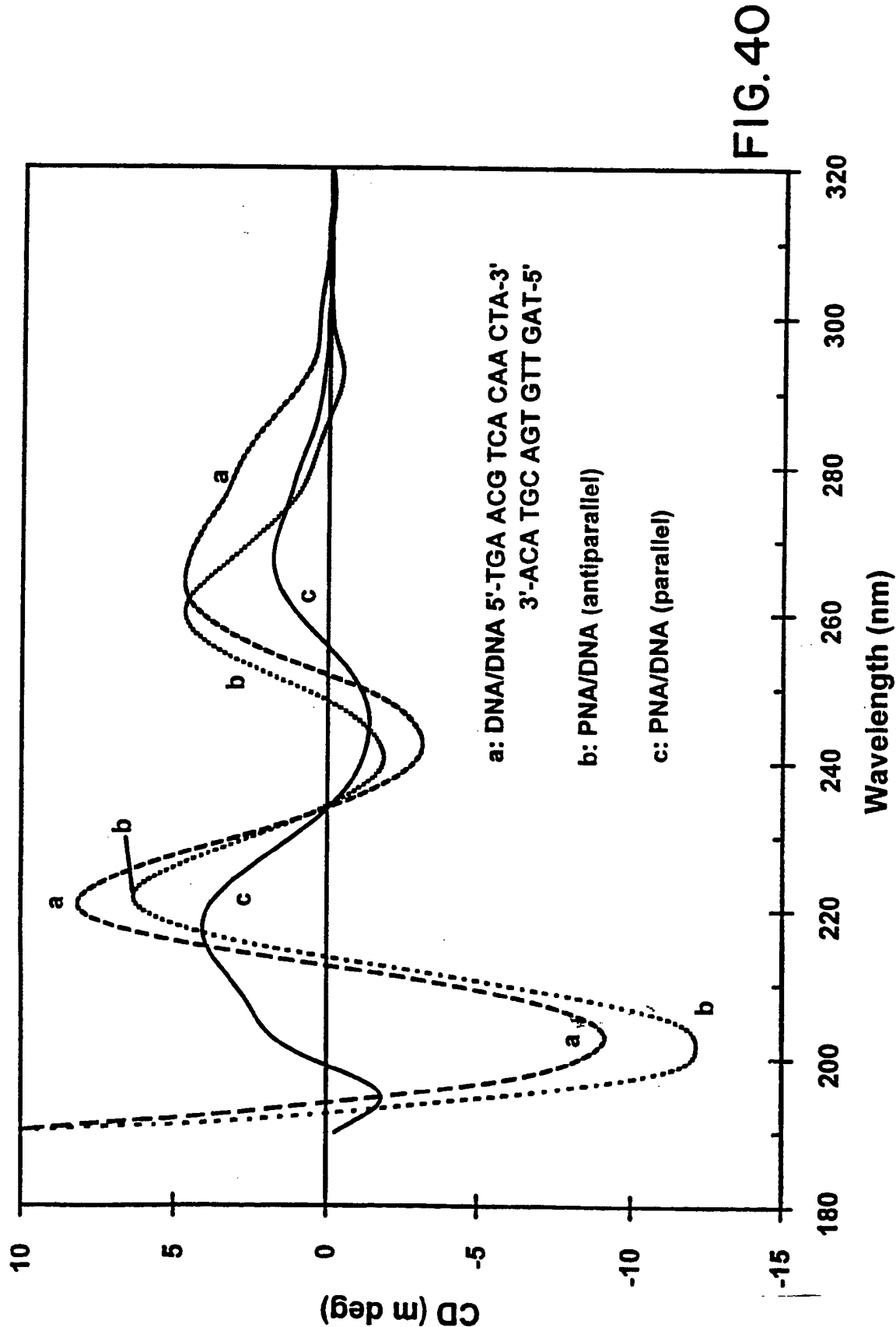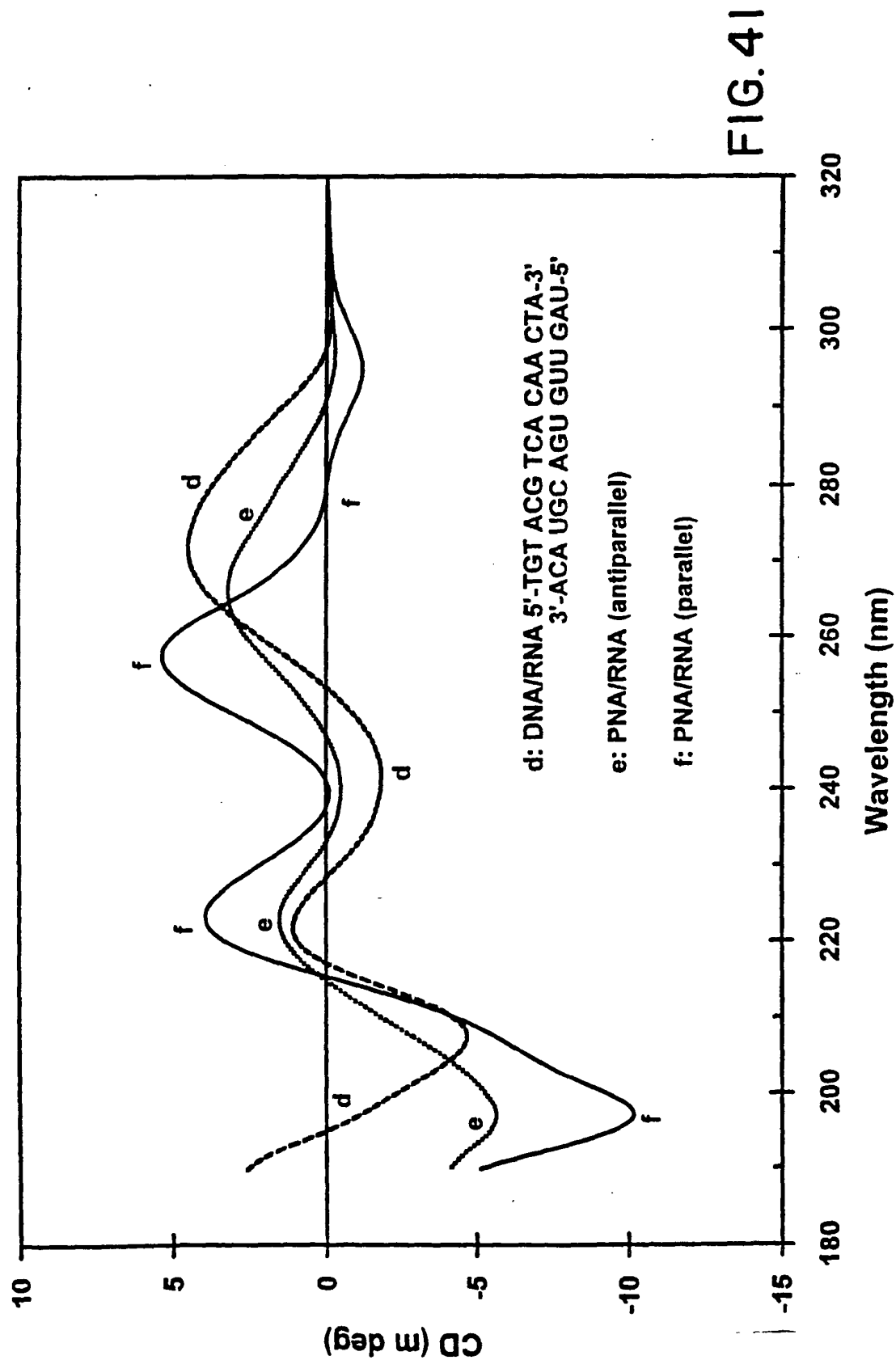
39 / 40



FIG. 42

SCHEME



# FIG. 43

# INTERNATIONAL SEARCH REPORT

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(5)    :C07K 5/00, 7/00; C12Q 1/68; C08L 77/00; A61K 48/00

US CL    :435/6; 514/44; 530/300, 350

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. :    435/6; 514/44; 530/300, 350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CA, WPI

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO, A, 92/20702 (BUCHARDT ET AL.) 26 NOVEMBER 1992, see especially page 2, lines 8-18; page 6, lines 17-27; page 7, lines 5-18 and pages 8-10. | 1-78 |

☐ Further documents are listed in the continuation of Box C.        ☐ See patent family annex.

* Special categories of cited documents:

"A"    document defining the general state of the art which is not considered to be of particular relevance

"E"    earlier document published on or after the international filing date

"L"    document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O"    document referring to an oral disclosure, use, exhibition or other means

"P"    document published prior to the international filing date but later than the priority date claimed

"T"    later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"    document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"    document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"    document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 02 OCTOBER 1994 | 0 7 OCT 1994 |

| Name and mailing address of the ISA/US | Authorized officer |
|---|---|
| Commissioner of Patents and Trademarks<br>Box PCT<br>Washington, D.C. 20231 | SCOTT HOUTTEMAN |
| Facsimile No.    (703) 305-3230 | Telephone No.    (703) 308-0196 |